ABDUCTION, DISPOSITIONS, AND ALTERNATIVES IN SCIENCE

On the Rational Reconstruction of Scientific Negotiation

Dissertation

zur Erlangung der Würde des Doktors der Philosophie

der Universität Hamburg

vorgelegt von

Alfred Nordmann

aus Bad Harzburg

Hamburg 1985

Ich warf allerlei Gedanken im Kopf herum bis endlich folgender obenhin zu liegen kam.

Georg Christoph Lichtenberg*

I verily belief free will & chance are synonymous. - Shake ten thousand grains of sand together & one will be uppermost, - so in thoughts, one will rise according to law.

Charles Darwin

(*I tossed all kinds of thoughts around in my head until finally the following one came to lie on top.)

Table of Contents

Introduction: Perspectives I

> Do you realize that what you are bringing up is the trick argument that a man cannot try to discover either what he knows or what he does not know? He would not seek what he knows, for since he knows it there is no need of the inquiry, nor what he does not know, for in that case he does not even know what he is to look for.

Socrates (in Guthrie's translation of Plato's *Meno*) can afford to call this paradox a 'trick argument'. He is satisfied that it does not seriously obscure the impulse to learn since his famous theory of *anamnesis* dissolves it at least in part: all learning is only re-cognition. And yet, he has only shifted the issue from epistemological to ontological grounds: now one has to discern the narrow confines of true re-cognition, and has to distinguish it from the mere illusion that one has actually succeeded in learning something. Thus, regardless of whether Socrates's elegant maneuver is adopted, all learned and learning scholarship has to intermittently reassure itself of its own sincerity in the face of the paradox: to what extent is the result of so-called learning nothing but a reflection of what one was looking for in the first place?

In this respect not unlike theories of art and literature, the science of science represents a consolidated interdisciplinary effort at *learning* from the workings and products of a socially instituted human activity. While the philosophy of science is one of the tributaries to that (as of yet) hardly autonomous discipline, learning from science is not and has not traditionally been the exclusive aim of the philosophy of science. In the tradition of natural philosophy, some philosophers of science interpret the world and relate on scientific principles the human being to the world, elaborating, in effect, a scientific *Weltbild*. And from a commitment to epistemology stems the ambition of others to guide and inspire, judge and advance the cause of science itself. In contrast, the following study belongs to that group of works which merely seeks to understand the institution, to learn in a philosophical as well as empirical manner about the ways in which the scientific community operates.

To ensure against the pitfalls and paradoxes of learning, a programme in the science of science may begin by analyzing into component parts the resolve to learn in any domain, including science. At the outset stands the conviction that there is something to be learned here: one starts on the presumption that somehow science makes a worthwhile subject of learning, for instance, that science represents a model of particularly efficient, successful, rational, and transparent collective reasoning. Secondly, since understanding science as a historical artifact requires a fair degree of descriptive accuracy, the intuitive preconceptions going along with the presumption of worthiness have to be tempered by cautious open-mindedness, sensitivity, if not humility. It seems thirdly that to strike the balance between preconception and caution, a plan of inquiry or principle of learning has to be adopted: while *some* unrevisable preconceptions are indispensable preconditions of learning, the explicit adoption of a plan puts them into a long-range perspective which may ultimately allow the recognition of their merits and defects. For instance, if that plan somewhat boldly assigns a firm role in the course of inquiry to the presumption of worthiness, intuitive preconceptions about science can be transformed into comparable hypotheses as variations on the adopted principle of learning are tried out. Thus, the presumption of worthiness with its preconceptions about the subject-matter is tempered by caution as a bold and tentative principle of learning is adopted. And yet, the impulse to learn is not aiming for tentative and cautious results: it is confidently based on the expectation that the presumption of worthiness pays off by providing insights which are generalizable to other domains of human activity and thus of universal value. Learning from and about science should be, as the famous dictum has it, learning for life.

I believe that the desire to learn (though not the process of learning) characteristically combines these four components. Awareness of them does not dissolve or avoid the paradox of learning so much as it provides a way of taking it into account. Philosophical temperament ultimately decides *how* this awareness should be incorporated within the inquiry into science and at what point one can be satisfied that the competing components are taken into account sufficiently well. As a matter of temperament, the ensueing investigation does not first resolve the friction between hard and fast preconceptions, humble caution, tentative boldness, and envisioned universal validity. Instead it tries to preserve that friction throughout the investigation. On the conviction that sensitivity to the deep problems underlying all inquiry should not stifle but inspire inquiry itself, the investigation sails forth with somewhat blue-eyed optimism into the domain of what (equally optimistically) is called the 'science of science'. The first chapter constructs a plan of inquiry by adopting a principle of learning. That plan emerges from a discussion of the competing demands for descriptive accuracy and normative validity on philosophical theories of science. In accordance with the plan of inquiry, the second chapter provides a conceptual framework which appropriately structures 'scientific activity', the subject-matter at hand. In the third chapter, the conceptual framework (a theory of alternatives in science) is projected upon a historical case. After the test-case has been thus prepared, the fourth and final chapter brings the plan of inquiry to bear on that episode in the history of science. Due to the foundational character of the argument, all this is much sketchier than it should be and hardly more than the outline of a comprehensive research programme in the science of science. Since that programme stands and falls with its usefulness for the inquiry into science, that sketchiness is surely both helpful and regrettable: helpful, if the programme's heuristic power can be judged by seeing whether overall a fortunate balance between caution and boldness has been struck, and regrettable as only meticulous study of various historical cases can bear out its heuristic power and allow for a judgement of its soundness. While the last two chapters discuss an episode from the history of biology (Darwin and 'Natural Selection'), more detailed further case-studies will have to supplement the present effort (e.g. on Newton and 'Universal Gravitation', or on Max Weber and the 'Spirit of Capitalism').

As a foundational reflection on the science of science, this study is most closely related to previous work by members of the so-called Starnberg-Group (Gernot Böhme, Wolfgang van den Daele, Wolfgang Krohn). To my mind, they originated the first research programme in the science of science which in a detailed, sophisticated, and (as of yet) not nearly sufficiently appreciated manner unites the history, philosophy, and sociology of science. Owing to this proximity, the relation between my enterprise and their theories is hardly documented in the course of the following

investigation itself. However, the four chapters amply testify to the influence of all the other works which have deeply impressed me in one way or another. As those works are customarily held to express the views of schools that are profoundly at odds with one another, my attempt to assemble Carnap and Popper, Kuhn and Lakatos, Peirce and Feyerabend, and many more under the umbrella of a unified programme in the science of science may not do justice to any of them. I hold accountable for this seemingly undisciplined eclecticism my own learning experience at Hamburg and Columbia Universities where I was exposed to teachers who themselves came from vastly different traditions and yet encouraged through their work (though all in their own styles) more integrative, non-factional, pluralistic perspectives. To them, I am most indebted. First and foremost, Lothar Schäfer and Robert K. Merton influenced me in many more ways than they can possibly imagine. Here I can thank them only for their encouragement, support, and almost uncanny confidence in my work. I am afraid that I have squandered their credit too liberally. Isaac Levi hopefully appreciates the transplantation of some of his thought into a line of reasoning which possibly appears quite alien to him. Either way, I am grateful for his extensive comments on early versions of chapters 2 and 4. Comments and suggestions by Kathrin Kaiser, Jörg Zimmermann, Wolfgang Detel, and Harriet Zuckerman also helped me at various stages of the work. I owe apologies to

two persons whom I am not giving enough in return for their inspiration: instead of a case-study on Darwin (which I promised Howard Gruber), there is only a brief discussion of a case-study which remains essentially unwritten; and that I do not live up to Harald Wohlrapp's expectations is plain by the lack of concreteness in thought and style which one finds so much in current philosophy of science and for which a more or less generous use of examples and illustrations is no substitute. Finally, I thank Janis Vieland for acting in critical distance as my philosophical as well as political conscience. If this work has any integrity at all, she shares all the credit for that.

A final note: The body of the text is strewn or littered with substantive footnotes - which, I hope, are insubstantial as far as the intelligibility of the main argument is concerned. They are generally tangential, anticipate upcoming results or remind the reader of previous phases in the argument, and they are often designed to clarify and avoid misunderstanding. Like all substantive footnotes, they have a jarring and complicating effect on the reading experience and one should perhaps consult them only as questions arise.

Rational Reconstruction and Theories of Science

According to Imre Lakatos, any philosophical theory of science provides the hard core for a research programme in the history of science. Given that the goal of such a research programme is to account in methodological terms for what scientists consider good scientific gambits[1], the competing theories can be compared and judged by their explanatory power. Similarly, one can treat theories of science as candidates for or test cases of normative theories of rationality. In this respect, they are primarily judged by their ability to withstand epistemological criticism. And thirdly, theories of science serve a social goal insofar as they contribute towards a justification of the (relative) autonomy of the scientific enterprise. In this context, they are best evaluated in terms of their expediency. These are three goals which philosophies of science can be expected to meet. There may well be more. It is readily apparent, however, that conflict between just these three goals could easily arise. For instance, Popper's methodology of critical rationalism performs very badly as a historiographical research programme while it provides a compelling epistemological theory which also offers a clear demarcation between science and any other social enterprise. The reverse holds true for Lakatos's methodology of scientific research programmes. While it presents a superior theory for the purpose of rational reconstruction in the history of science, it represents a weak test case for any theory of rationality and does not contribute much to the problem of how to distinguish science from non-scientific activities[2]. Which theory is to be preferred?

For the resolution of such conflict and thus for the possibility of adjudicating once and for all the merits of competing theories of science, a number of conditions would have to be satisfied. That is, a meta-theory is needed which lists and ranks all of the relevant goals to be satisfied by a theory of science. But such a meta-theory can resolve the conflict between, say, Popper's and Lakatos's methodologies only if it also provides operational criteria by which one can determine to what extent a given theory does indeed meet any such goal. None of these conditions are satisfied as of yet and it seems fair to say that no such meta-theory is in sight. None will be provided here. The remainder of this chapter exemplifies rather than elaborates or derives a tentative strategy concerning this meta-theoretical matter. That strategy really consists of hardly more than the resolve to be sensitive to the issue of the competing goals. In the context of a largely historiographic enterprise such as this, that resolve amounts to this: the theory of science which will be developed here is to be kept open to future qualifications arising from the fulfillment of the other two *desiderata*. More concretely, while the theory of science that will best meet the goal of the historiographical research programme will not be very restrictive, it has to allow for an ensueing trade-off between the merits of inclusiveness and the application of minimal restrictions that will distinguish it socially and epistemologically as a theory specifically of science. To this end, the following argument aims at establishing a methodological framework that is compatible with any more restrictive theory of science. The issue of rationality in science will then be raised in a way that preserves this compatibility: by representing various theories of science as providing optional methodological postures which scientists can choose to adopt under various circumstances.

i.

The problem of competing cognitive goals that a theory of science might be expected to meet has been circumvented for now. But Lakatos's contention that methodologies function as research programmes in the history of science poses rather complex problems of its own. The following discussion of these problems will lead to the formulation of some general constraints on philosophical theories of science in respect to rational reconstructions of the history of science. While disagreeing with its particular conclusions, it attempts to take seriously and to preserve the intuitions underlying Lakatos's article "History of Science and its Rational Reconstructions":

> the methodology of historiographical research programmes implies a pluralistic system of authority, partly because the wisdom of the scientific jury and its case law has not been, and cannot be, fully articulated by the philosopher's statute law, and partly because the philosopher's statute law may occasionally be right when the scientists' judgment fails. I disagree, therefore, both with those philosophers of science who have taken it for granted that general scientific standards are immutable and reason can recognise them *a priori*, and with those who have thought that the light of reason illuminates only particular cases. The methodology of historiographical research programmes specifies ways both for the philosopher of science to learn from the historian of science and *vice versa*.[3]

The main argument of Lakatos's article consists of two parts. First, he shows that if one takes falsificationism as a meta-criterion for appraising historical research programmes, one will find any such programme falsified[4]. This is true, regardless of whether the historical research programme rests on a falsificationist, inductivist, or conventionalist methodology, or on Lakatos's own methodology. This finding is well in tune with the passage just cited: the relationship between history and methodology is simply such, that history 'falsifies' any methodology. But what if one uses Lakatos's methodology of research programmes as a meta-criterion? In the second part of his argument, Lakatos shows that on this criterion history corroborates any theory of science – to varying degrees, however[5]. The one theory to be chosen is then the one that is *best* corroborated by the history of science. "In the light of better rational reconstructions of science one can always reconstruct more of actual great science as rational."[6] On these terms, Popper's falsificationism outperforms inductivism and Lakatos's methodology of research programmes fares still better.

The two parts of Lakatos's argument show that a methodology of research programmes is to be preferred over falsificationism as a historiographical meta-criterion, but no more: the failure of falsificationism as a historiographical meta-criterion does not entail anything about its performance as a methodology for (natural) science. On the contrary, *that* failure can be predicted by falsificationism itself by contrasting historiogrpahy and (natural) science in precisely this respect: historiography of science is not properly a scientific discipline, and in Popperian terms that assertion is equivalent to saying that falsifications do not play a decisive role in the historian's work[7]. And concerning the second part of Lakatos's argument as presented so far, a falsificationist is not at all compelled to accept that a rational reconstruction is *better* if it rationally reconstructs *more* of science. In view of the normative character of falsifiability as a criterion, falsificationists do not have to rationally reconstruct anything but what according to their methodology qualifies as rational scientific behavior in the first place[8].

Lakatos does not let the falsificationist get away quite so easily. The question is not just *how much* of science can be rationally reconstructed, but rather how much of *great* science. By introducing the notion of 'great science', Lakatos appeals to the intuition that much of science *is* rational – an intuition that is entirely independent of any *a priori* criteria demarcating rational from irrational, scientific from non-scientific behavior. That is, scientific rationality is a historical given. Therefore, normative theories of scientific rationality should also be expected (though they cannot be forced) to conform to this antecedent intuition.

> After all, one must admit (*pace* Popper) that until now all the 'laws' proposed by the apriorist philosophers of science have turned out to be wrong in the verdicts of the best scientists. Up to the present day it has been the scientific standards, as applied 'instinctively' by the scientific *elite* in *particular* cases, which have constituted the main – although not the exclusive – yardstick of the philosopher's *universal* laws.[7]

Thus, history provides an explanandum which is sufficiently well described as 'great science', 'the verdicts of the best scientists', and 'the scientific standards as applied instinctively by the scientific *elite*'. Philosophers of science are not compelled to devise an explanans for this explanandum, but so much the better if they do[10]. The argument against apriorism is no more and no less than a persuasive appeal. Lakatos can conclude only on the basis of this appeal that the methodology of research programmes allows for the better rational reconstruction of history and that therefore it is to be preferred as a theory of science.

At this point, however, a somewhat frivolous suggestion may be introduced. It comes from quite a different - namely the 'anti-theoretical' - camp[11]: once it is acknowledged that scientific rationality is independently given and, so to speak, embodied in 'great science' and the verdicts of the scientific elite, why bother to construct a methodology at all which posits what scientific rationality is? Thomas Kuhn exemplifies the attitude underlying this suggestion. He makes this point in his response to Lakatos's article on history and its rational reconstructions. Kuhn is startled that Lakatos imputes to him the heretical view that science is irrational, while he, Kuhn, finds Lakatos's work so congenial to his own. Attempting to resolve this asymmetry, Kuhn identifies a sentiment shared by him and Lakatos:

> Scientific behavior, taken as a whole, is the best example we have of rationality. Our view of what is to be rational depends in significant ways, though of course not exclusively, on what we take to be the essential aspects of scientific behavior. That is not to say that any scientist behaves rationally at all times, or even that many behave rationally very much of the time. What it does assert is that, if history or any other empirical discipline leads us to believe that the development of science depends essentially on behavior that we have previously thought to be irrational, then we should conclude not that science is irrational but that our notion of rationality needs adjustment here and there.[12]

On the basis of this shared intuition concerning the rationality of science, Kuhn (as opposed to Lakatos) does not feel compelled to deal with the issue of rationality. His historical theses simply *cannot* seriously call into question scientific rationality. They will at most challenge some untenable limitations imposed by philosophical notions of rationality. The scientist is not obliged to live up to a philosophical idea of rationality. On the contrary, it is the philosopher's responsibility to keep up with the scientist's behavior[13]. In an article about causality in social science, Alasdair MacIntyre elucidates the approach which - on this view - a philosopher of science would have to adopt: following Aristotle, action is to be considered as the conclusion of a practical syllogism.

> The action [...] follows from the premises in just the way in which a proposition follows from the premises of a theoretical syllogism. We can bring out the force of the 'follows from' in the last sentence by considering a close parallel with the theoretical syllogism. This is the case when a

> speaker appears to accept the premises of a valid syllogism and yet to resist the conclusion. He allows that a warm front brings rain and also that a warm front is approaching, but he refuses to say 'So it's going to rain'. There are only two possibilities open in such a case. *Either* the speaker thinks that an additional premise needs to be added to the argument, *or* he does not realize to what he has committed himself. If the latter is ruled out, then we can infer what kind of additional beliefs the speaker has, e.g. that he thinks the warm front may well be intercepted in its course by some other meteorologically relevant factor. In precisely the same way, where a speaker affirms premises and fails to act on them, we can make inferences as to his additional beliefs. [...] it has therefore emerged that to look for the antecedents of an action is not to search for an invariant causal connexion, but to look for the available alternatives and to ask why the agent actualized one rather than another. [...] The explanation of a choice between alternatives is a matter of making clear what the agent's criterion was and why he made use of this criterion rather than another and to explain why the use of this criterion appears rational to those who invoke it. In other words, the internal relation of beliefs to action is such that in explaining the rules and conventions to which action in a given social order conform we cannot omit reference to the rationality or otherwise of those rules and conventions. Explaining actions is explaining choices; and explaining choices is exhibiting why certain criteria define rational behaviour for a given society. To this we must now add that the beginning of an explanation of why certain criteria are taken to be rational in some societies is that they *are* rational.[14]

Lakatos considers the anti-theoretical approach as the extreme antithesis to apriorism – and equally undesirable at that. Where apriorism was not informed at all by the history of science, the anti-theoretical stance appears not at all informed by epistemology. Both run counter to his intuition that the two modes of inquiry into science should complement each other[15]. The following two objections to the anti-theoretical approach can be identified in Lakatos's argument:

> A) If methodologies stop providing normative standards of rationality, the ability to systematically compare them relative to scientific practice will get lost.
>
> B) If not complemented by normative considerations, the anti-theoretical approach leads to an undue generalization that implies counter-intuitive results.

As each of these two objections is scrutinized quite extensively, the contours of a conceptual framework for a theory of science provided by scientific practice will emerge as a significant modification of Lakatos's methodology of research programmes.

In regard to objection A), the disagreement between Lakatos and Kuhn does not concern the role of rationality in the scientific enterprise, but Kuhn's apparent renunciation of the desire to systematically investigate that role[16]. Lakatos argues that comparing theories of science in terms of their historiographical power hinges on the comparability of the historiographical research programmes relative to a partially standardized explanandum. That is, it hinges on the explanatory relation of normative theories of rationality to behavior that is (antecedently) identified as rational. Given that a historical event can be described in many ways and that explanation and prediction pertain to events under particular descriptions, a comparison of predictive or explanatory success presupposes a certain degree of uniformity in the description of events across the various theoretical approaches. This view rests on two separate suppositions, both requiring justification:

(a) Uniformity in the description of the events can be established by taking the judgement of the scientific community on what is or was good and rational science for the purpose of determining, focussing, or standardizing the explanandum.

(b) This uniformity is not only a necessary, but also a sufficient condition for ensuring meaningful comparability of various theories of science in terms of their explanatory power.

While Lakatos has an (implicit) argument for (a), he apparently takes (b) for granted – thus jeopardizing his entire project.

Objection A) to the anti-theoretical approach has to be scrutinized first in respect to supposition a). A suitable starting-point is Lakatos's characterization of his theory of science in terms of its explanatory power:

> Where Kuhn and Feyerabend see irrational change, I predict that the historian will be able to show that there has been rational change. The methodology of research programmes thus predicts (or, if you wish 'postdicts') novel historical facts, unexpected in the light of extant (internal of external) historiographies [17]

Surely, for an historical fact to be novel it need not designate an event which is discovered only now. A novel fact may well be an event under a new or different description. Take this event:

> (1) 1800, Priestley publishes *The Doctrine of Phlogiston established and that of the Composition of Water refuted.*

Let it also be known as context for (1) that a far-reaching consensus concerning the rejection of phlogiston theory and the acceptance of Lavoisier's theory (maintaining *inter alia* the composition of water) emerged in the late 1780s. Further, that Priestley"s book did nothing to erode this consensus[18]. The event described in (1) therefore signifies a particularly tenacious hold on a (for all practical purposes) obsolete theory. Accordingly, it can be redescribed as

> (1a) 1800, Priestley expresses his tenacious hold on phlogiston theory by publishing *The Doctrine of Phlogiston established and that of the Composition of Water refuted.*

Both, (1) and (1a) state common historical knowledge. And yet, the words chosen in (1a) are arbitrary in view of a multitude of legitimate redescriptions of (1). Consider

> (1b) 1800, Priestley expresses his tenacious hold on a (for all practical purposes) obsolete theory.
>
> (1c) 1800, Priestley does not acknowledge the falsification of phlogiston theory.

(1d) 1800, Priestley reminds the scientific community of some unresolved questions concerning the conflict of the two chemistries.

(1e) 1800, Priestley still believes that phlogiston theory is true.

(1f) 1800, on the subject of chemistry Priestley reasserts his francophobia.

One can easily imagine explanatory theories for any one of these redescriptions of event (1). A historian working in Popper's spirit will not consider Priestley's tenacity as properly scientific and might provide an explanation of (1c) in terms of extrascientific factors[19]. Another historian (in the tradition, say, of Bernal) will take (1f) as a typical expression of the interplay between science and ideology, and will explain (1) as the erection of a scientific superstructure on the base of national and class interests. The historian who turns to (1d) will view Priestley's action as an ordinary contribution in the course of the communal search for certified public knowledge. Those three historians have entirely incompatible views on science, but their theories cannot be compared in terms of whether they explain (1). For, *any* historiographic research programme can explain (1) under *some* redescription. And each redescription of (1) is a novel (but not unexpected or forbidden) fact relative to all those other research programmes which do not explain *it*. In reference to (1), no research programme can thus be said to explain more than any other. If the prediction of novel facts is to serve as a criterion for comparing rivaling theories of science, some sort of *ceteris paribus* condition is needed in order to establish a standard of comparison. Lakatos himself provides one when he states his criteria for 'falsification':

> For the sophisticated falsificationist a scientific theory *T* is *falsified* if and only if another theory *T'* has been proposed with the following characteristics: (1) *T'* has excess empirical content over *T*: that is, it predicts *novel* facts, that is, facts improbable in the light of, or even forbidden by *T*; (2) *T'* explains the previous success of *T*, that is, all the unrefuted content of *T* is included (within the limits of observational error) in the content of *T'*; and (3) some of the excess content of *T'* is corroborated.[20]

It is condition (2) which functions as a *ceteris paribus* condition here: everything else being equal, the prediction of a (corroborated) novel fact by *T'* makes *T'* preferable over *T*. That is, the prediction of a novel fact helps decide between two theories if in principle both are answerable to the prediction of that particular fact under roughly the same description. (2) is designed to ensure that both *T* and *T'* explain subsets *S* and *S'* of more or less the same class of facts such that *S* is contained in *S'*[21]. Lakatos's approach thus requires that the class of facts under consideration can be identified by some shared characteristic. For this purpose, he rather ingeniously suggests that the predicate 'rational' should invariably be added to the explanandum. Accordingly one arrives at *inter* *alia*

> (1') 1800, Priestley publishes *The Doctrine of Phlogiston established and that of the Composition of Water refuted*: a rational act.
>
> (1a') 1800, Priestley expresses his *rationally* tenacious hold on phlogiston theory by publishing his *The Doctrine of Phlogiston established and that of the Composition of Water refuted.*
>
> (1c') 1800, Priestley *rationally* does not acknowledge the falsification of phlogiston theory.
>
> (1f') 1800, on the subject of chemistry Priestley *rationally* reasserts his francophobia.[22]

There are still any number of redescriptions for (1'), all of them containing the predicate 'rational'. But they have been standardized in such a way that the proposed explanation have to be of quite a different character than before. All of the explananda are now focussed towards that predicate and each now requires an appropriately focussed explanans. Using an example of Arthur Danto's, one might say that adding 'rational' to the above listed redescriptions of (1) has very much the same effect as adding 'unpopular' to the explanandum 'Louis XIV died':

> you cannot say that Louis XIV died unpopular because he ate poisoned lobster: for *that* only explains why he *died.* Indeed, it would not even enter into an explanation of why he dies unpopular.[23]

The partial standardization of the explananda pays off right away. Any redescription of (1') is a novel and unexpected, indeed forbidden, fact for Popperians[24]. The redescriptions of (1') therefore have less variability (and are less arbitrary) than the redescriptions of (1). The redescriptions of (1') all belong to 'more or less the same class of facts', namely the class of rational actions. While Popperians had no trouble explaining (1c), all redescriptions of (1') (and particularly (1c')) present anomalous facts in the context of their inquiry[25]. As long as a non-apriorist stance is maintained (as argued above), a theory that explains (1') is to be preferred over Popperian falsificationism.

Lakatos defines as 'internal history' those historical accounts which are based on this standardization. Any theory that yields more internal history of science is to be preferred over one that yields less. Whoever abandons (or omits) a normative characterization of 'rationality' is out of this race from the start. The verdicts of the scientific community are needed to standardize the explananda, and the normative aspects of a theory of science are needed to generate explanations for these explananda. The class of explananda is now given and so are theories of science. And one can now determine whether a given theory does or does not generate explanations for the explanandum.

So far, Lakatos's argument is compellingly clear. But now the question arises whether the ability to determine the explanatory success of competing theories of science suffices for an evaluation of these theories so that Lakatos's intuitions are fully satisfied. Lakatos seems to think that it does. But closer scrutiny of this second supposition (see above, page 11) shows that it does not.

To explain an event is to subsume it under a general law[26]. Consider the following three scenarios:

(2a) From a conjunction of general laws from various domains (e.g. social psychology and decision theory) the following general law has been deduced:

(x)(t) (If x is faced with a theory opposing x's own theory, and if that opposing theory originated in a country which has undergone a political revolution that challenges the political structure of x's country; if x is considered a religious non-conformist in x's own country, and if x is involved in a priority dispute with a supporter of the opposing theory over an experimental discovery which served as crucial evidential support to that theory: then it is rational for x to tenaciously hold on and defend x's own theory regardless of whether it has been rejected by the international scientific community)

Together with the appropriate initial conditions, this general law could form an explanans for a redescription of (1'). It should be noted that its extension is logically unlimited, but that – as a matter of fact – the history of science provides only one instance where the conjunction of the initial condition actually obtains, i.e. the case of Priestley[27].

(2b) A redescription of (1') is explained in two steps, i.e. only after a redescription of (1) has been explained:

1st step:

L: (x)(t)(... => ...)
C: [initial conditions]

E: [a redescription of (1)]

2nd step:

L*: (x)(t) (If x is a scientist and x performs an action in his or her role as a scientist, then the action of x is rational)
C*: Priestley is a scientist and in this role performed action E [a redescription of (1), the explanandum of the 1st step]

E*: Priestley rationally performed action E [a redescription of (1')]

Here it is L* which really does the job – regardless of what kind of general claim is made by L.

(2c) Since none of the familiar philosophies of science predict (or postdict) Priestley's action, a fictional methodology is assumed for this third scenario which in some aspects resembles Lakatos's and Feyerabend's proposals. Based on Lakatos's conviction that any research programme may always stage a comeback, it predicts that there always will be scientists who engage in the rational strategy of counterinduction aimed at the exploitation of just this possibility of a comeback[28]:

(x)(Ey)(t) (If x is a theory or research programme and x has recently been replaced by another theory or research programme, then y will rationally uphold x and do so publicly)

This general law will serve as the major premise in an explanans with a redescription of (1') as its explanandum.

The explanatory power of the three theories of science which generated the preceding explanations, cannot be compared in respect to (1'). The best of the three will be the one that rationally reconstructs most episodes in the history of science aside from (1'). If one wants to find it, the theories involved in the three scenarios ought first to be identified. The conjunction of general laws in (2a) probably presents only a small subset of the methodological propositions generated by its underlying theory of science. Indeed, it may well be impossible to list the entire set of propositions which constitutes that methodology. There are no such problems with (2b). L* is a blanket-law and captures in a rather elementary way the 'anti-theoretical' posture which Lakatos ascribes to Kuhn: L* is *all* the methodology there is[29]. Only scenario (2c) may have been generated by what traditionally and intuitively is considered a proper theory of science. As each of these theories succeeds in designating *this* bit of history as internal history of science, how do they fare in the overall rational reconstruction of the history of science? As it stands, no matter how strong the theories underlying (2a) and (2c) are, obviously they can *at best* achieve what (2b) or an equivalently non-committal theory achieves: the weakest, nearly vacuous and irrefutable methodology makes the most general claim and therefore has the greatest explanatory power. It explains *any* action as rational which is designated as a scientific action.

The only way of avoiding this result is to separate acceptable from unacceptable explanatory schemes, e.g. to stance finally comes into play (see above, page 10). The scope of that presumption will have to be clarified, and a further constraint on theories of science in respect to rational reconstructions will be introduced.

The rejection of apriorism was based on the intuition that rationality is embodied in (great) science and therefore to be accounted for in historiographic terms. This argument was subsequently generalized so as to state: since science *is* rational, there is no need to theorize about the conditions under which science is a rational enterprise. L* in (2b) was taken to be a paradigmatic statement of the anti-theoretical stance. It maintained that any action performed by scientists in their roles as scientists is a rational action. Now, Lakatos's second objection to the anti-theoretical generalization consists in pointing out that it leads to counter-intuitive results: L* is simply false. Indeed, both Kuhn and MacIntyre do acknowledge that a vast generalization like L* cannot be sustained. Concerning Priestley, Thomas Kuhn writes:

> Though the historian can always find men – Priestley, for instance – who were unreasonable to resist for as long as they did, he will not find a point at which resistance becomes illogical or unscientific. At most he may wish to say that the man who continues to resist after the whole profession has been converted has *ipso facto* ceased to be a scientist.[31]

First of all, this passage confirms what was said before:

(philosophical) theories of science in respect to rational reconstructions of the history of science. Following upon the disdain for apriorism, a historiographical approach has to be developed which - for the sake of explanatory power - is as much as possible like the one utilized in scenario (2b). However, this inclination towards the anti-theoretical approach should lead to the construction of a conceptual *framework* for a methodology that is subsequently provided by scientific practice itself. As a historiographical research programme its promise or interest would lie not so much in its broad and near-trivial explanatory power. After all, that explanatory power would seem to follow simply from the restraint from providing a proper methodology that would include a normative conception of 'truth', 'objectivity', and 'rationality' in science. Within that conceptual framework, the history of science should be representable as a succession of decisions or choices on (limited) options, where the absence of such proper methodology functions as a presumption of rationality in science, and consequently as a heuristic commitment to the elucidation of that presumed rationality in scientific decision-making on a case by case basis. Thus, the promise or interest of the present approach would lie in this step by step construal of current methodologies at various stages in the history of science. However, as that presumption of rationality reaches its limit, Lakatos's second objection to the anti-theoretical devise rules according to which somehow scenario (2b) somehow does not count as a legitimate contender. A falsifiability-requirement would be one such rule. But specifying the rules for distinguishing acceptable from unacceptable explanatory schemes is tantamount to formulating a (methodological) theory of science: after all, a methodology provides criteria for evaluating explanations and therewith guides the search for acceptable explanations. The strategy for avoiding the unwelcome result thus leads straight into a (vicious) circle: one cannot very well presuppose a theory of science as one selects a criterion for the selection of a theory of science.

So, once the unity of the subject matter has been established by partially standardizing the explanandum, and once it is assured that the explanatory schemes which are to be compared actually pertain to the same sort of thing, one is faced with all the problems that are traditionally associated with the notion of 'explanation'. To be sure, these problems can be approached in an apriorist fashion by a normative philosophy of science which does not aspire to generate a most successful historiographic research programme for rational reconstruction in the history of science. But this is precisely how the problem arose in the first place: from not wanting to be apriorist[30].

This leads to the adoption of a first constraint on either one chooses an apriorist approach and *rules* that Priestley was not a scientist anymore, or one assumes an anti-theoretical point of view, adding maybe an inconsequential *proviso* such as a general acknowledgment that of course not *everything* scientists do is really rational. Such inconsequential *provisos* were already contained in the more general formulations of the anti-theoretical approach that were cited above:

> That is not to say that any scientist behaves rationally at all times, or even that many behave rationally very much of the time.[32]

Kuhn states the exception but does not act on it, i.e. does not explore this distinction. Similarly, MacIntyre simply rules out one of two possible interpretations:

> *Either* the speaker thinks that an additional premise needs to be added to the argument, *or* he does not realize to what he has committed himself. If the latter is ruled out, then we can infer what kind of additional beliefs the speaker has [...][33]

According to Kuhn there are, and according to MacIntyre there may be instances which are not justifiably designated as rational by a general law of type of L*. In order to construe these as falsificatory instances, one has to assume that not all that is designated as 'scientific activity' is rational[34]. However, since the methodology of research programmes was adopted as the meta-criterion for the historiography of science, the existence of such falsifications would not warrant the rejection of L* in absence of a better rival. Thus still within the framework of L*, Kuhn's and MacIntyre's *provisos* raise the following question: Is there a way of distinguishing between good and bad scientific gambits without reverting to apriorism?

ii.

There is a fairly simple and straighforward answer to this question. It leads to the adoption of a principle of investigation which shall guide the construction of a theory of science that supersedes Lakatos's historiographic methodology.

Is there a way of distinguishing between good and bad scientific gambits without reverting to apriorism? Yes, scientists do it all the time. However, as they do so, their concern is not the rationaliy of scientists but the expediency of certain choices towards the achievement of certain goals. Their judgements on what is good for science and scientific progress, serve in effect to differentiate good from bad scientific gambits in such a way that the issue of rationality does not even come into play. Accordingly, philosophers of science should limit themselves to pluralistically identifying the processes involved in making these judgements in the course of scientific practice. By identifying rather than postulating methodological practices, they will be able to explore the middle ground between apriorism and the anti-theoretical approach from the scientist's point of view.

It will be remembered that Lakatos himself aimed for that middle ground between apriorism and the anti-theoretical approach. He expressed that by saying that the history and philosophy of science should mutually inform each other. So far, however, the critical discussion of Lakatos's intuitions has led to the point where one can seee that obviously the history of science can inform philosophy. But since the weakest philosophical theory will have most explanatory power, it remains quite unclear how philosophy is to inform science or the historiography of science. Some of the middle ground can be regained by outlining a programme for approaching the issue of scientific rationality without formulating a normative methodology, while preserving one's intuitions concerning 'rationality' as well as the inclusiveness of the anti-theoretical position. On the assumption that rationality is not at all different in science that in any other context of decision-making, the forthcoming approach shall therefore be subsumed under a general theory of decision-making. For historiographic purposes, this amounts to an acceptance of the view that decisions reveal (rational) preference, i.e.

that from the outcome of a decision one can conjecturally reconstruct it as rational relative to appropriate information, options, and goals. The presumption of rationality thus functions as a heuristic. It is used to understand the decision-making processes of scientists themselves and leads to the formulation of hypotheses as to what information, and which options and goals must have been available in any given case[35].

Framing the issue of rationality within general decision-theory is not to deny that the scientific enterprise is uniquely distinguishable from any other sort of social enterprise. On the contrary, specifically 'scientific' features can be discovered in the situation in which scientific decisions have to be made. These features may consist of constraints on the formulation of the alternative options on which scientists can deliberate, on the (professed or real) goals which the choice among the options has to be related to, and on the arguments or rules which can be used to relate the chosen alternative to that goal. And it is precisely here where traditional methodologies and other normative theories of scientific conduct enter into the process of scientific negotiation: for instance, they may be used to characterize the options which confront scientists in their decision-making, especially insofar as those options pertain to 'explanation' and 'prediction'. Also, if specifically scientific decision-making is so constrained that scientists may publicly assume only acceptable or even certified[36] methodological postures in order to argue for their choice, then traditional methodologies enter into the theory of science as suppliers for the argumentation repertoires of scientists.

The second constraint on (philosophical) theories of science in respect to rational reconstructions in the history of science has thus emerged: philosophy should inform the historiography of science by framing the problem of rationality within general decision-theory. Further, such philosophical problem-areas as 'explanation' and 'prediction' should be framed as scientific issues which are negotiable by scientists themselves. That is, a theory of alternatives in science is to be culled from philosophical reflections on scientific method. This second constraint supplements the first one insofar as it helps to taxonomically differentiate processes in the history of science which are left completely undifferentiated by the adoption merely of blanket-law L*. The adoption of these two constraints designates a problem-shift in the philosophy of science: instead of deliberating on problems concerning the rationality of science or scientists, philosophers are now asked to structure the problems upon which scientists deliberate themselves.

The approach proposed here is clearly distinct from one that formulates a methodological theory of science. For, it maintains only that general decision theory, a theory of alternatives in science, and the set of (at any given time) acceptable methodological postures, provide together a repertoire that is sufficient for the business of rational reconstruction[37]. The shift of focus towards constrained options and methodological postures resulted from the realization that, left alone, the presumption of rationality cannot be usefully employed: it either trivializes an inquiry into science since it will demonstrate that just about any action is rational, or it has to be fortified by aprioristic considerations which are undesirable for the historiographic reasons that were discussed above. This is how the observation that scientists themselves show no concern with the issue of rationality but a great concern with what is (properly) scientific, provides a master-clue for this inquiry. And therefore, as scientific activity is characterized by certain constraints on options and methodological postures, the presumption of rationality has to be extended to include the presumption of scientific conduct. By the same token, the notion of 'rational reconstruction' is narrowed so that it includes only rational reconstruction with reference to acceptable methodological postures. That is, the rationality of scientists as indivdual agents cannot be jeopardized by a theory of science. However, historical inquiry has to probe the extent to which scientists can defend their positions within the constraints of scientific argumentation. A modified version of L* should therefore rest on this presumption of rationality and scientific conduct. Though still very weak, this version of L* is not trivially satisfied anymore and rational reconstruction can fail on occasion. However, how easily it can fail depends on the constraints on methodological postures which the scientific community has imposed on itself at any given time.

Lakatos's methodology of scientific research programmes has now been challenged by a sketch for a pluralistic theory of science. That sketch preserves Lakatos's basic intuitions:

- the history of science should inform the philosophy of science, and *vice versa*,
- the notion of 'rationality' provides the link between history and philosophy of science since it gives rise to the programme of 'rational reconstruction' in the history of science,
- a philosophical theory of science is more progressive (in respect to extant theories of science) if it

> enables the historian of science to reconstruct more of the history of science as rational.

But the pursuit of these intuitions has led to more radical conclusions than those presented by Lakatos himself - conclusions which were subsequently relativized to standards of acceptability as embodied in scientifc practice itself. The result can be summarized in two ways. Firstly, by reconsidering Lakatos's definition of 'internal history' as 'rational history in the light of some normative theory of rationality'. Since it was shown that just about any decision is rational in light of an appropriately weak normative theory of rationality, this definition has to be amended in accordance with the second constraint on theories of science: 'internal history' is that history which is shown to be rational in reference to admissable scientific arguments. And secondly, this result can be summarized by stating a deliberately fuzzy principle of investigation (PI) as a constraint on theories of science in respect to rational reconstructions:

> (PI) A theory of science has to provide a minimally restrictive normative theory of rationality (e.g. general decision-theory presuming rationality along the lines of L*), and minimal restrictions on the system of alternative options, acceptable arguments, and goals. These minimal restrictions are thought to be specified and supplemented in all kinds of ways in the course of scientific negotiation.

A theory of science in accordance with (PI) deserves the label 'pluralistic' because of the words 'in all kinds of ways'. Though it sounds exceedingly fuzzy, this phrase is a far cry from the infamous 'anything goes', insofar as, in principle, it can be rendered far more precise for any particular episode in the history of science, and since in any such episode it may then present only an extremely limited range of admissable arguments to the working scientist. In this, the pluralism envisioned here differs from other pluralistic theories of science[38]. It does not first establish a pluralism of methods as a preferred strategy for the working scientist, and thus only *a fortiori* for the scientist of science: it does not matter on this account whether or not it so happens that working scientists dogmatically adhere to one set of methodological rules[39]. Scientists of science should always anticipate that there is a plurality of methods which is manifest in the scope of scientifically admissable and even philosophically certified arguments at any given time, and which is now available to them for the purposes of rational reconstruction. Of course, if scientists of science discover regularities in the process of scientific negotiation which obtain at all times and which go beyond what is asserted by a theory of science that satisfies principle (PI), the theory of science can be strengthened and improved on the basis of empirical results. In this respect, principle (PI) inaugurates self-correcting empirical research-programmes in the historiography of science[40].

Aside from general decision theory, a theory of alternatives in science and an investigation into the scope and use of argumentation repertoires will have to embellish principle (PI) for a full-fledged theory of science to emerge. But nothing has been said and little need be said about the notion of scientists' 'goals': there are no antecedent reasons to assume that the adoption of certain goals further scientific progress while others impede it. Regardless of whether scientists wish to find truth for its own sake, to enhance the glory of God, to establish 'thoroughgoing materialism', to undermine the present social order, or to prove the inferiority of women to men - as long as they produce scientifically acceptable arguments that speak to the matter of a currently pressing alternative in science, the ultimate goals which are to be furthered by these arguments need not be taken into consideration[41]. This is both to admit the tremendous strategic influence of external forces on the development of science and to discount their importance as one deliberately narrows the focus on the forms of conduct which socially and epistemologically differentiate the scientific enterprise from decision-making in other spheres of life[42].

iii.

The relation of internal to external forces on the development of science provides a suitable testing-ground on which the proposal that was developed here can be shown to constitute a progressive problem-shift *vis-a-vis* Lakatos's methodology of scientific research programmes. It will be recalled that Lakatos defines 'internal history' as 'rational history in light of some normative methodology', and that this crucial intuitive link between theories of science and historiography was amended only slightly so as to read 'rational history in light of some (weak) normative methodology in reference exclusively to scientifically acceptable arguments'. The task of the philosophy of science consists in showing *more* of the history of science as internal history. Thomas Kuhn has pointed out that Lakatos's conception of 'internal history' deviates somewhat from that of the historian.

> In standard usage among historians, internal history is the sort that focuses primarily or exclusively on the professional activities of the members of a particular scientific community: What theories do they hold? What experiments do they perform? How do the two interact to produce novelty? External history, on the other hand, considers the relations between such scientific communities and the larger culture. [...] Obviously there is much overlap between normal usage and Lakatos'. [...] Lakatos' internal history is far narrower than that of the

> historian. It excludes, for example, all consideration of personal ideosyncrasy, whatever its role may have been in the choice of a theory, the creative act which produced it, or the form of the product which resulted.[43]

The tension between Lakatos's narrower and the historian's wider conceptions may be a temporary one. If philosophy of science is to show *more* of science as internal, then some day a sufficiently advanced philosophical theory of science may be able to show as rational just what historians already call 'internal' science. On this view, the borders of 'internal' vs. 'external' are fluid. After all, Lakatos directly links the notion of 'internal science' to the relative progressiveness of a given philosophical theory of science:

> When a better rationality theory is produced, internal history may expand and reclaim ground from external history.[44]

For Lakatos and his methodology of scientific research programmes, however, the tension still exists and the historian's usage is still wider than his - as he himself points out in his discussion of priority disputes in science.

> According to Merton "scientific *knowledge* is not the richer or the poorer for having credit given where credit is due: it is the social *institution* of science and individual men of science that would suffer from repeated failures to allocate credit justly". But Merton overdoes his point: in important cases (like in some of Galileo's priority fights) there was more at stake than institutional interests: the problem was whether the Copernican research programme was progressive

> or not. (Of course, not all priority disputes have scientific relevance. For instance, the priority dispute between Adams and Leverrier about who was first to discover Neptune had no such relevance: whoever discovered it, the discovery strengthened the same (Newtonian) programme. In such cases Merton's external explanation may well be true.)[45]

Here Lakatos identifies an episode in the history of science which cannot be rationally reconstructed with his methodology of scientific research programmes. And accordingly, he considers Merton's explanation external. However, one could well argue that that explanation "focuses primarily or exclusively on the professional activities of the members of a particular scientific community", and that it should thus be considered an internal explanation in accordance with the historian's common usage. After all, even where priority disputes become dysfunctional ends in themselves which do nothing to advance the growth of knowledge or to support the reward system in science, they are still altogether internal to the fabric and social structure of science: a dysfunctional priority dispute is a self-regulating device of the scientific community gone mad. Robert Merton expresses this point:

> It has sometimes been said that the emphasis upon recognition of priority has the function of motivating scientists to make discoveries. [...] From this, it would seem that the institutional emphasis is maintained with an eye to its functional utility. [...] Once it becomes established, forces of rivalrous interaction lead it to get out of hand. [...] Rationalized as a means of providing incentives for original work and as expressing esteem for those who have done

> much to advance science, it becomes transformed into an end-in-itself. It becomes stepped up to a dysfunctional extreme far beyond the limit of utility.[46]

But regardless of whether pre-theoretically one wishes to consider the priority dispute between Adam and Leverrier as internal or external, the theory proposed here is able to reconstruct it rationally as internal history while Lakatos cannot. Indeed, any involvement in a priority dispute by a scientist can be shown to be rational and scientific, especially when the scientist invokes arguments which implicitly or explicitly appeal to, say, the institutional importance within the social structure of science that credit be given where credit is due. By counting appeals to the professional norms of the scientific community as an acceptable methodological posture, this proposal gives a more progressive theoretical account of the internal-external relation than does Lakatos's methodology. And since, furthermore, the set of acceptable methodological postures is determined by scientific practice rather than by methodological verdicts, there is no antecedent limit to how much ground internal history is yet to reclaim from external history.

iv.

The proposal sketched out here can now be evaluated in respect to the various goals which a philosophical theory of science can be expected to meet. Firstly, the conceptual framework for a methodoloy as provided by general decision theory, a theory of alternatives in science, and a repertoir of methodological postures constitutes at best a weak and minimally restrictive normative theory of rationality. Since on the view expounded here, scientific decision-making is not different in kind from any other form of decision-making, nothing about rationality is to be learned by looking at the scientific community in particular[47]. Secondly, the goal of providing a strong historiographical research programme can be met by the proposed approach. This cannot be surprising, though, since this is what the proposal is designed to provide[48]. And lastly, as a social theory serving the justification of the (relative) autonomy of science as an institution, the proposal does not by itself lead to policy statements, such as whether or not it is desirable to preserve that institution as it now exists. It does suggest, however, to what extent contemporary science is characteristically defined by internal processes feeding on the social discreteness not of the goals of scientists, but of the other elements going into scientific decision-making. The very notion of a collection of methodological postures that scientists are free to assume presupposes a social institution which supports this kind of autonomy of choice. The conditions under which the autonomous institutional structure is maintained are seen as part of what scientists are concerned with in their roles as professionals (e.g. in their rational decisions to adopt or reject a proposed theory), which further ties social theory of science to the methodological proposal made here.

This proposal in relation to other philosophies of science has been labeled above as 'pluralistic'. Weak and minimally restrictive as it may be, the conceptual framework, together with next chapter's theory of alternatives in science, captures however little consensus there is among philosophers of science. Equipped with principle (PI) one can proceed from this smallest common denominator by recognizing the validity of various philosophies of science as bonified strategies towards the rational resolution of scientific problems. As strategies or elements of the overall argumentation repertoire, however, each of them is just one among many.

As far as there is any 'deep' philosophical conviction involved in this project, it could finally be stated like this: rationality consists in the ability to correctly adduce reasons to one's decisions to act or not to act, to believe or not to believe in certain ways. From the point of view of a theory of rationality, the formal correctness is all that matters (consistency and coherence). In the spirit of the Enlightenment, 'rationality' is not perceived as an issue of philosophical interest but rather as a fundamental if not unalienable property of the human mind, of human thought. Whether the reasons themselves are good or bad, appropriate or inappropriate is not a question of rationality, it can only be settled within a particular context of inquiry. Simplifying greatly, one might say that a metaphysical approach to science establishes how scientific reasoning arrives at the truth. Methodological orientations are divorced from this metaphysical question and try to specify good principles of reasoning in scientific inquiry. This proposal finally suggests (and it is not the first to do so[49]) that one should go yet one step further, and leave the reasoning to the scientists, in order to focus only on whatever may constrain it and how these constraints are made to serve the growth of objective knowledge[50].

Notes to Chapter 1

[1] The label 'theory of science' is used to designate any philosophical, sociological, or historical theory of science. 'Methodology', on the other hand, designates a more or less elaborate system of rules (or certifiably admissable arguments) which is generated by a (philosophical) theory of science.

[2] An argument for these assessments of Popper's and Lakatos's work will emerge in the course of this chapter.

[3] Lakatos [1978], p. 137.

[4] Lakatos [1978], pp. 123-131.

[5] Lakatos [1978], pp. 131-136.

[6] Lakatos [1978], p. 132.

[7] The contrast between science and the presumably non-scientific discipline 'history of science' figures prominently in the work of Popper's student, the historian of science Joseph Agassi. Unfortunately, his formulations on the subject are somewhat obscure. It seems that he ascribes the impossibility of a scientific history of science to a structural feature of historical subject matter itself, i.e. to the feature that historical agents have a consciousness which differs (and has to differ) from that of the historian and which cannot be treated by the historian as causally effective:

> the testing ground should, of course, be history: a reformer cannot, but an observer should apply his theory to history. It turns out that things are not so simple, and for the following, rather obvious reason. The history of science is largely the history of what we value in science, and so the reformer can and indeed should rewrite history. [...] And so we need to know the difference between the discoverer's reconstruction and the legislator's reconstruction - which leaves us exactly where we were before. How can the reformer apply his ideas, say, to Newton, if Newton knew nothing about these ideas? [...] When we study physics we [...] must fix today's physics - at least long enough to give it a good look, to teach it to our young rascals, etc. But we need not banish history. [...] But when we treat the living actions of scientific

> inquiry we hypostatize not only today's method, which is inevitable, but also yesterday's, which is unnecessary. Newton, we say, preached inductivism but practiced my methodology. [...] Clearly, both Bacon and Descartes said, a scientist must be aware of his method. Must. Duhem and Popper say, he need not be, and is, in fact, seldom aware of what he is doing. Query: according to Duhem and Popper, is his awareness a factor significant to his research in any way whatsoever? Suppose we say, no. Then, first and foremost, all arguments from the way people happen to have looked at science become strictly external. (Agassi [1981], pp. 227 and 235ff.)

Having thus shown the impossibility in principle of (purely) internal history of science (history of science rationally reconstructed by e.g. Popperian means), he salvages the historian's work: by not making a claim to scientific knowledge (in a Popperian sense), historians cannot be held accountable to such a claim:

> you have a theory about what is internal history and how it works; you apply the theory; it applies with difficulty; you introduce external factors to overcome the difficulty. Query: is this move quite kosher? or should your theory be rejected? The answer hinges on the question, is your theory introduced as purely internalist? The answer to this hinges on, can there ever be a purely internalist theory? If we say, there cannot be a purely internalist history, then your rescue operation looks more kosher than if we say, there can be a purely internalist history. (Agassi [1981], p. 64)

To be sure, Agassi's problem does not arise in the Lakatosian framework. In contrast to Agassi who rescinds the claim that there can be (purely) internal history, Lakatos modifies the notion that the difficulties of applying a theory to history should therefore lead to its rejection even if no better theory (edging closer to purely internal history) has been proposed. The issue whether there can be *purely* internal history stems from a Popperian juxtaposition of theory (internal) and falsificatory instances (external). Lakatos's 'sophisticated falsificationism' requires only that a better theory of science produces *more* internal history in respect to extant theories. Thus, as opposed to Lakatos and on the basis of a largely Popperian commitment, Agassi denies the possibility of a science of science which is self-exemplifying in respect to its conception of 'scientific method'. -- Popper himself has rather more principled reasons for not taking recourse to the history of science in the attempt at gaining an understanding of science:

> how can the regress to these often spurious sciences help us in this particular difficulty? Is it not sociological (or psychological, or historical) *science* to which you want to appeal in order to decide what amounts to the question 'What is *science*?' (Popper [1972], p. 58)

Thus, while Lakatos thinks it desirable that the science of science should be self-exemplifying as far as its conception of 'science' is concerned, Popper shies away from the specter of an infinite regress.

8 There is nothing viciously circular about this, considering that this theory was not originally designed to meet the goal which Lakatos ascribes to all theories of science, i.e. the goal of providing an explanatory theory for the history of science. On the contrary, Popper's is a paradigm example of a normative theory of rationality in the guise of a theory of science.

9 Lakatos [1978], pp. 136f.

10 The use of 'explanans' and 'explanandum' underlines the continuity between science and Lakatos's conception of the science of science (compare above, note 7): ideally, the science of science should explain scientific activity (in history and in the presence) methodologically in very much the same way as natural scientists explain natural phenomena. -- As Mittelstrass [1974] points out, however, the historian's work cannot be completely reduced to this conception of a science of science:

> Causal explanations of past events are not [...] examples of *historical comprehension* (Kausale Erklärungen für zurückliegende Entwicklungen sind [...] keine Beispiele für *historisches Begreifen*. (p. 122)

Whatever is meant here by 'historical comprehension', adopting this view is not tantamount to denying that historical comprehension consists largely of the ability to adduce (causal) explanations.

11 Lakatos [1978], pp. 136f. -- With the juxtaposition of 'apriorism' with the 'anti-theoretical' stance, Lakatos's a bit heavy-handed, if not derogatory labels are reluctantly adopted here.

12 Kuhn [1971], p. 144.

[13] Kuhn stated this position nowhere quite as clearly as in the passage which was quoted just now. However, clear evidence for his attitude can also be found in Kuhn [1972], p. 264 where he evaluates an argument by Feyerabend and himself on the role of reason in science:

> to describe the argument as a defence of irrationality in science seems to me not only absurd but vaguely obscene. I would describe it [...] as an attempt to show that existing theories of rationality are not quite right and that we must readjust or change them to explain why science works as it does. To suppose [...] that we possess citeria of rationality which are independent of or understanding of the essentials of the scientific process is to open the door to cloud-cuckoo land.

Insisting that scientists are rational while hesitating to codify this aspect of scientific behavior results in severing 'truth' or 'proof' as intra-theoretical notions from inter-theoretical criteria employed in the (rational) discourse on 'better theories'. Compare p. 266 of the same article, and p. 263:

> My argument [...] emphasizes that, unlike most disciplines, the responsibility for applying shared scientific values, must be left to the specialists' group. It may not even be extended to all scientists, much less to all educated laymen, much less to the mob.

For a more sophisticated statement of this programme see also Kuhn [1983].

[14] MacIntyre [1962], pp. 53 and 61.

[15] The conflict between the anti-theoretical and the apriorist attitudes attests to the presence of a problem in the philosophy of science which is common to all attempts at learning from something: the problem of balancing one's responsiveness to historical contingency with one's preconception that there is something non-contingent about the structure of science which makes it a worthwhile subject of investigation.

[16] Kuhn's renunciation does not involve the denial that there are worthwhile problems associated with the investigation of the role of reason in science. On the contrary, he puts considerable effort into arguing that he does not wish to trivialize or render redundant issues concerning truth and justification. Yet he insists that there are modes of theorizing about science (such as the one employed by himself) which are indifferent to any outcome of research into those deep issues. Compare e.g. Kuhn [1983], pp. 569f.:

> the problem of induction, [...] viewed from the perspective developed here, acknowledges that we have no rational alternative to learning from experience, and asks why that should be the case. It asks, that is, not for a justification of learning from experience, but for an explanation of the viability of the whole language game that involves 'induction' and underpins the form of life we live. To that question I attempt no answer, but I would like one.

Kuhn segregates two levels of analysis. The level of analysis on which he himself is at home avoids this issue of justification. Leaving it at that would amount (in the words of Mittelstrass [1974], p. 123)

> *to discharging the analytical conception of science from the obligation to comprehend the history of the sciences in a justificatory manner as the history of a process of mediation between theory and* praxis (*Entlastung eines analytischen Wissenschaftsverständnisses von dem Erfordernis, die Geschichte der Wissenschaften in begründungsorientierter Weise als die Geschichte einer Vermittlung von Theorie und Praxis zu begreifen*)

But surely, from a given research perspective it is quite sufficient to do what Kuhn did, namely to acknowledge that there is a further justificatory level of analysis which one may not be equipped to address quite yet.

17 Lakatos [1978], p. 133. - Lakatos considers 'prediction' as roughly equivalent with 'stating what is (or was) to be expected'. The same can be said of 'explanation', which provided the basis for Hempel's 'symmetry of explanation and prediction'. It is thus appropriate to discuss Lakatos's historiographical problem as a problem of 'explanation' in terms of the hypothetico-deductive model of explanation. (On the malleability of the H-D model with respect to greatly different views on the character of historical explanation see Danto [1965], pp. 201-232. - Compare also note 10 above.)

18 Aside from these few matters of historical record, this appeal to Priestley is not based on an inquiry into Priestley's work and life.

19 Compare Lakatos [1978c], pp. 202f.:

> A catastrophical consequence of a narrow methodology is that, as well as impoverishing actual problem-situations, it invokes *external* - psychological, sociological - explanations

> because its *internal* framework of rational explanation fails too soon. [...] For instance, the Popperian insistence on abandoning a theory after the 'crucial experiment', opens the door to those trendy 'sociologists of knowledge' who are trying to explain the further - and possibly unsuccessful - development of the rival programme as the irrational, wicked, reactionary obstinacy of established authority against enlightened revolutionary innovation.

20 Lakatos [1978b], p. 32.

21 "More or less the same class of facts" and "roughly the same description" are about as precise as Lakatos's own "explaining the previous success of *T*". This seems to be an ineliminable lack of precision: if an event is explained only under a description relative to the explanans, then (whenever *T'* contains theoretical terms defined differently than those of *T*) *T'* might in some sense explain the success of *T* but will not explain events under the same description as *T* did - it will not explain precisely the same facts. The paradigm example for this is the progressive problem shift of Einstein's theory in relation to Newton's. This lack of precision may not be partially damaging to Lakatos's account. Either way, nothing much hinges upon it here.

22 It is not quite clear whether this represents an adequate reading of Laktos's article. Conceivably, he intended to add 'scientifically rational' rather than 'rational' to each explanandum. However, this would render his argument unintelligible. After all, it would presuppose an independently given notion of 'scientific rationality' in contradistinction to plain or ordinary 'rationality' - and Lakatos gives no indication where that notion could be taken from. As interpreted here, 'rational' in the explanandum means only 'rational in the light of any explanatory theory of science'. A privileged kind of 'scientific rationality' which is somehow discontinuous with 'rationality' need not be assumed.

23 Danto [1965], p. 310.

24 There is another advantage resulting from Lakatos's move: the arbitrariness in redescribing (1) is substantially diminished as one selects (1') as the event to be redescribed: that is, the attribute 'rational' is introduced as the verdict of the scientific community. Relative to the central predicate 'rational', the gulf between 'raw event' and redescription is now largely eliminated: with (1'), the event "comes" under a

description that is provided by the scientific community. This is how the history and practice of science informs the theory of science.

25 It should be noted that (1f'), for instance, is still a novel (and neither unexpected nor forbidden) fact relative to any historical research programme which is not out to explain *it*. As with the various redescriptions of (1), this is of no importance and does not enter into the question as to whether some theory performs better than any of the others: the relevant difference is that any theory of science explains (1), while only some can explain (1') (not, for instance, Popperian theory of science).

26 Since a more extensive account of 'explanation' will follow in the next chapter, discussion of whether this brief statement is satisfactory and how it should be amended will be postponed until then. Suffice it to say that the reader is free to pass judgement on the following three 'explanations'. Each of them embodies its own standard of what might count as an acceptable 'explanation' and at least two of the three appear blatantly inadequate. Either way, the argument to be presented here is independent of whether some of these 'explanations' are inadequate or not.

27 A somewhat similar scenario was proposed as a criticism of Lakatos by Hall [1971], pp. 157f.

28 Compare for instance Lakatos [1978], p. 114:

> what for the falsificationist looks like the (regrettably frequent) phenomenon of irrational adherence to a 'refuted' or an inconsistent theory and which he therefore relegates to *external* history, may well be explained in terms of my methodology *internally* as a rational defence of a promising research programme.

29 Of course, Kuhn does not go quite so far. It is not easy to precisely determine his position. From the ensueing criticism of L* may emerge a closer approximation to his actual view.

30 What leads Lakatos to a disdain of apriorism forces him to acknowledge the tremendous historiographical power inherent in the anti-theoretical approach. It has been pointed out repeatedly that Lakatos (probably for that reason) falls short of providing a (normative)

methodology himself. See for instance Schäfer [1974], pp. 217-221. – Lakatos is aware of this and presents it as a

> shift in the problem of normative philosophy of science. The term 'normative' no longer means rules for arriving at solutions, but merely directions for the appraisal of solutions already there. (Lakatos [1978], p. 103)

31 Kuhn [1970], p. 159. – Compare this to Lakatos [1978], p. 117:

> One may rationally stick to a degenerating programme until it is overtaken by a rival *and even after*. What one must *not* do is to deny its poor public record.

Concerning Priestley, then, firm anti-anti-theorist Lakatos may turn out to be more liberal than anti-theorist Kuhn.

32 Kuhn [1971], p. 144 – for more of the context, see above page 8.

33 MacIntyre [1962], p. 53 – for more of the context, see pages 8f. above.

34 Some such construal is necessary also for a somewhat strengthened version of L*:

> (x)(t) (If x is a scientist and x performs an action in his or her role as a scientist which the scientific community does not openly condemn as improper, then the action of x is rational)

The absence of open condemnation is taken as implicitly certifying the rationality of x's behavior. Here one has to assume that intuitions concerning rationality can at times override the tacit verdict of the community. Thus, Lakatos's disagreement with the proponents of laws such as L* has nothing to do with the extent to which the verdicts of the scientific community are incorporated into the formulation of a general law.

35 Whether this presumption can be shown to hold for all cases or not is a sceondary issue: having adopted it as a heuristic, the burden of proof lies in the empirical work of the decision-theorist or of the historian aided by decison-theory.

36 Acceptable, that is, to the scientific community – and possibly certified by the scrutiny of logico-epistemological criticism. Inductivism, for

instance, is a certified methodological posture. (It will be presented as such in the final chapter.)

[37] The remainder of this study is devoted to an elaboration and exploration of this approach. However, no attempt will be made to delimit the set of acceptable methodological postures, and no theory provided on how they may change over time. Further, it will remain quite open whether there is some algorithm according to which only certain decisions are rational given certain options. Finally, the notion of 'general decision theory' will not be elaborated in detail.

[38] Compare e.g. Feyerabend [1978], Naess [1972]. The notion of pluralism that is employed here is closer to Schäfer [1974] or Merton [1981].

[39] Radnitzky [1971] presents theoretical and methodological monism and pluralism as postures which may be characteristic of certain phases in the history of science. In order to be sensitive to this characteristic, the historian and philosopher of science has to anticipate theoretical and methodological pluralism.

[40] The research programmes generated by (PI) should be designated as 'sensitive' rather than 'frigid'. These categories were introduced by Knorr-Cetina [1981], pp. 17ff. as she presented her ethnographic approach to scientific reasoning. That approach led her to the laboratory as a strategic research-site in the science of science. However, her sensitive approach needs to be supplemented by inquiry into scientific argumentation in formal contexts, that is, in contexts where the constraints on acceptable methodological postures are enforced. An understanding of science and the production of scientific knowledge has to pay attention to the contexts of discovery (in the laboratory) *and* justification (constrained negotiation). The line between sensitive and frigid approaches does not coincide (as Knorr-Cetina suggests) with the line dividing the investigation of methodological issues form the ethnographer's immersion in the laboratory. (The exchange with Knorr-Cetina will be continued in further footnotes: so far, a claim has been made; a full argument in its support will be developed only at the end of the last chapter.)

[41] One might say that acceptable scientific arguments are all geared towards the (professed) ultimate goal of 'truth' or 'fit of theory to reality'. That is, the constraint on the set of admissable argument is tantamount to a constraint on the publicly statable

goals: by staying within the realm of scientific argumentation, scientists act *as if* they were interested only in truth. As far as an analysis of the (institutional) structure of science is concerned, it really does not make a lot of difference whether or not the professed goals always coincide with the personal of individual scientists. Habermas's 'theory of communicative action' lays some groundwork for a research programme that would explore scientific action in respect to this simultaneous pursuit of goals. Especially pertinent is his discussion of the performative features of utterances and argumentation, where (following Austin [1967], pp. 94-131) he distinguishes between perlocutionary and illocutionary acts that are simultaneously performed by utterances. Compare Habermas [1981], pp. 389-392:

> For illocutionary acts the *meaning of what was said* is constitutive, and likewise, for teleological acts the *intentions* of the agent. One now gets what Austin calls *perlocutive effects* as illocutionary acts take on a role within a teleological context of action. [...] From these considerations Austin has drawn the conclusion that illocutionary success stands in a conventional or internal relation to the speech-act, while perlocutionary effects remain external to the meaning of what was said. (Wie für illokutionäre Akte die *Bedeutung des Gesagten* konstitutiv ist, so für teleologische Handlungen die *Intention* des Handelnden. Was Austin *perlokutive Effekte* nennt, entsteht nun dadurch, dass illokutionäre Akte eine Rolle in einem teleologischen Handlunsgzusammenhang übernehmen. [...] Austin hat aus Überlegungen dieser Art die Konsequenz gezogen, dass illokutionäre Erfolge mit der Sprechhandlung in einem *konventionell* geregelten oder *internen* Zusammenhang stehen, während perlokutionäre Effekte der Bedeutung des Gesagten äusserlich bleiben.)

For the purpose of understanding (socially situated) communication, one has to rely on the illocutionary character of utterances (p. 394). Thus, traditional analyses of the aims or goals of science should be interpreted as attempts at making explicit the illocutionary force of acceptable scientific arguments. These analyses provide answers to the question: why does it make sense for scientists to adopt (and enforce) communicative constraints (as e.g. on methodological postures)?

42 This (as well as the preceding note) indicates that,

as far as the historian's work is concerned, rational reconstruction covers only a small though significant slice of scientific activity. By taking into consideration what is insignificant for the purpose of rational reconstruction, the historian creates an interplay of internal and external histories of science. A drastic example illustrates this point: *X* is a scientist who is known to be a misogynist. Whenever a woman produces work relevant to his scientific specialty, he will subject her contribution to scrutiny far exceeding the scrutiny he applies to work by his male colleagues. Intending to destroy her career, he proceeds to publish only what is most damaging to her findings. Owing to his socialization as a scientist and the peer-review system in science, his published article will be free from polemics and contain no statement of his misogynist position. Here, two histories have to be written just as two different courses of action will have to be taken against *X*: the arguments produced by him will have to be taken on their own terms regardless of the judgement on the person and the integrity of his motives.

[43] Kuhn [1971], p. 140.

[44] Lakatos [1978], p. 134.

[45] Lakatos [1978], p. 116.

[46] Merton [1973], pp. 321f. – In a footnote Merton highlights the notion of "stepping up patterns to unanticipated extremities".

[47] What may be said, however, is something like this: decision-making in science tends to be more scrupulous, more transparent than decision-making in other spheres of life. This, of course, makes the scientific community a good research site for anybody interested in rationality. But that doesn't make science or scientists more or less rational than other decision-makers. (It is assumed here that a normative theory of rationality is 'weak' if it is applicable to ordinary, everyday decision-making as much as to scientific decision-making. This use of 'weak' is not to pre-empt or trivialize attempts at formulating the conditions which make everyday decision-making rational.)

[48] Whether or to what extent historians should use it, is quite a different issue – just as the question to what extent historians are or should be involved in seeking out or corroborating general laws or theories of history. The question posed and answered here is merely: what kind of philosophical theory of science *would* and *could* perform better or worse as a successful historiographical

research programme.

49 Within the Popperian or deductivist tradition, Lakatos unwittingly and Feyerabend quite deliberately have moved in the same direction. Isaac Levi's epistemology, which stands in a Carnapian or inductivist tradition, also captures this intuition quite substantially.

50 To be sure, a metaphysical interest in science is not obsolete. Any theory of 'truth' or 'objectivity' will introduce further normative constraints on the notion of 'rationality' (in the sense that it is rational to follow true principles). These theories are particular interpretations of the formal, reductionist model of 'rationality' that is provided by general decision-theory. And these interpretations become manifest as optional methodological postures. They are therefore not appropriately considered as providing final words on the enterprise of science. Compare Schäfer [1974], p. 46:

> The history of science teaches us that the task of philosophy - if it has a task - cannot be one of complementing, i.e. completing science, but that it consists rather in the defence of its openness and autonomy. (Die Wissenschaftsgeschichte lehrt uns, dass die Aufgabe der Philosophie, wenn sie eine hat, nicht die einer Ergänzung, d.h. Abschliessung sein kann, sondern eher in der Verteidigung ihrer Offenheit und Autonomie besteht.)

A Theory of Alternatives in Science

As a first step towards an account of scientific decision-making, a theory of alternatives in science needs to be developed. The goal of such a theory is to give a somewhat precise characterization of the alternatives and options scientists deliberate and act upon. From this follows almost trivially the constraint that it should somehow designate specifically *scientific* alternatives. Therefore, the theory cannot be constructed without *some* antecedent notion of what science is. Here, however, a second constraint enters in. According to the considerations of the previous chapter, such an antecedent notion should be weak and unrestrictive; i.e., a theory of alternatives in science should be specific, maximally inclusive and minimally pre-emptive[1].

Balancing these demands is the constructive task before us[2], and it involves a series of steps.

- Firstly, a weak characterization of 'scientific activity' has to be given. Any number of approaches could be employed here. I propose that such a characterization can be constructed by somewhat eclectically combining the classical explication of 'explanation' by Hempel and Oppenheim with the criterion of empirical significance proposed by Carnap in *Testability and Meaning*[3]. A diachronic dimension is introduced by projecting Peirce's theory of abduction, deduction, and induction as the three stages in the development of a theory onto the largely synchronic treatment of scientific activity by Carnap, Hempel and Oppenheim.

- Secondly, a precise typology of alternatives has to be culled from this characterization. To be sure, alternatives can be expected to arise due to the very weakness of the characterization which does not uniquely determine the outcomes of many decisions scientists might face. The challenge here consists merely in finding a sufficiently precise and unified way of framing the typology.

- Finally, I simply propose that a 'specifically scientific alternative' is any alternative arising from 'scientific activity'. Thus, whether an alternative is specifically scientific or not is determined solely in terms of its genesis from scientific activity.[4]

Of these three steps, the last one does not warrant detailed scrutiny. The remainder of this chapter is therefore devoted to the problems of characterizing scientific activity and of finding a typology of alternatives arising from this characterization.

i.

What is 'scientific activity'? Very loosely, one can say that scientists deal with perceived regularities or structural properties of nature and that they are concerned with establishing relations between such regularities for the purposes of explanation and prediction. Other goals often attributed to science, such as 'finding the truth' or 'yielding technological control', seem largely dependent upon explanation and prediction[5]. Matters become even simpler as we adopt the view that explanation and prediction share the same logical structure: propositions used for the purpose of explanation can or could also be used for the purpose of prediction or *vice versa*[6]. This view was developed by Carl G. Hempel and Paul Oppenheim in 1948. It states the following four conditions for 'true scientific explanation':

> (R1) The explanandum must be a logical consequence of the explanans; in other words, the explanandum must be logically deducible from the information contained in the explanans [...]
>
> (R2) The explanans must contain general laws, and these must actually be required for the derivation of the explanandum. [...]

(R3) The explanans must have empirical content; i.e., it must be capable, at least in principle, of test by experiment or observation. [...]

(R4) The sentences constituting the explanans must be true.[7]

These four conditions provide a broad, primarily synchronic characterization of scientific activity. They constitute the deductive-nomological model of scientific explanation. By modifiying (R1) and substituting a measure of inductive support (like 'makes practically certain') for strict deducibility, one can accomodate inductive-statistical explanations in science as a variant of that model. Speaking of both model and variant, Hempel emphasizes that they

> are not meant to describe how working scientists actually formulate their explanatory accounts. Their purpose is rather to indicate in reasonably precise terms the logical structure and the rationale of various ways in which empirical science answers explanation-seeking why-questions.[8]

Accordingly, the four conditions are here considered as a conventional device for representing explanations[9]. The main virtue of this device lies in its very broad applicability. To be sure, Hempel cautions that his models of scientific explanation

> Obviously [...] are not intended to reflect the various senses of 'explain' that are involved when we speak of explaining rules of a contest, explaining the meaning of a cuneiform inscription or of a complex legal clause or of a passage in a symbolist poem, explaining how to bake Sacher

torte or how to repair a radio.[10]

But he *can* claim that the deductive-nomological model permits a conjectural formal reconstruction of all contexts in which the word 'because' occurs. A largely terminological consideration supports this claim. Initially, Hempel and Oppenheim proposed (R1)-(R4) in order to explicate the structure of 'scientific explanation'. Conditions (R1)-(R3) were presented as 'logical conditions of adequacy', (R4) as the 'empirical condition of adequacy'. The distinction was subsequently underlined in a footnote which was later added by Hempel.

> Requirement (R4) characterizes what might be called a correct or *true explanation*. In an analysis of the logical structure of explanatory arguments, therefore, that requirement may be disregarded.[11]

Thus, whatever fulfills conditions (R1)-(R3) may be called a 'scientific explanation', regardless of whether, ultimately, it is shown to be true or false. Now, just as (R4) selects a subset of 'true explanations' among all 'scientific explanations', (R3) selects a subset of 'scientific explanations' among all 'explanations'[12]: anything conforming to (R1) and (R2), i.e. any proposition that can be represented as an abbreviation for a set of propositions in deductive-nomological form is an 'explanation' - maybe a trivial, empty, meaningless and false explanation, but an explanation nonetheless. And if (R1) and (R2) can be taken as the core of the deductive-nomological model, then, indeed, any occurence of the word 'because' can be represented in that model, if only due to the heuristic power provided by (R1) and (R2). Thus, if someone says "x is late because she got held up in traffic", we can satisfy (R2) by inventing an appropriate general law ("whenever someone makes an appointment without anticipating a delay and gets held up in traffic, that person will be late"). That general law, in conjunction with some initial conditions ("she made an appointment, didn't anticipate a delay, and got held up in traffic") deductively entails the explanandum ("she is late"), thus satisfying (R1). Now, whether the person who just used the word 'because' meant to commit herself to the general law or not: we may say that by using 'because' she had to commit herself to some such general law[13]. In less trivial fashion, Hempel himself has shown in much of his writing that historical, teleological, functional and statistical explanation as well as explanation in reference to intentions can be represented using first of all (R1) and (R2) and only then (R3)[14].

If (R1) and (R2) ensure broad *prima facie* applicability of the H-O model as a conventional device for the representation of any explanation[15], (R3) and (R4) serve to delimit *scientific* explanation - hopefully without introducing undue constraints upon scientific activity. As we turn to these latter conditions, we ask in effect, how an explanation that is only a candidate for 'scientific explanation' can *become* and be recognized as scientific, and then how *it*, only a candidate for 'true (scientific) explanation' can *become* just that. Thus, the very distinction between 'explanation', 'scientific explanation', and 'true explanation' implies a temporal sequence, expressing the sequential formation of subsets or the consecutive order of decisions that have to be made once an explanation has been proposed. This diachronic dimension to the H-O model[16] can be highlighted using the later Peirce's theory of research, by correlating conditions (R1) and (R2) to the abductive phase in the development of a theory, (R3) to deduction, and (R4) to induction[17]. Accordingly, the latter two conditions will be considered within this temporal framework. According to Peirce,

> Abduction is the process of forming an explanatory hypothesis. [...] Its only justification is that from its suggestion deduction can draw a prediction which can be tested by induction, and that, if we are ever to learn anything or to understand phenomena at all, it must be by abduction that this is to be brought about.[18]

"Abduction merely suggests that something *may be*"[19]. It proposes an explanation, but only a candidate for 'true scientific explanation'. As such, it has to be presentable in syllogistic form as required by (R1) and (R2).

> [...] abduction, although it is very little hampered by logical rules, nevertheless is logical inference, asserting its conclusion only problematically or conjecturally, it is true, but nevertheless having a perfectly definite logical form. Long before I first classed abduction as an inference it was recognized by logicians that the operation of adopting an explanatory hypothesis - which is just what abduction is - was subject to certain conditions. Namely, the hypothesis cannot be admitted, even as a hypothesis, unless it be supposed that it would account for the facts or some of them. The form of inference therefore is this:
> The surprising fact, C, is observed;
> But if A were true, C would be a matter of course,
> Hence there is reason to suspect that A is true.[20]

Conditions (R1) and (R2) explicate the form of inference according to which C could be "a matter of course" merely by virtue of A's being true: A has to be some sort of general law in conjunction with assumed instantiations of the antecedents of that law, i.e. in conjunction with initial conditions; and C (the explanandum) has to deductively follow from A (the explanans)[21]. Thus, once an explanation is proposed for a fact or a class of facts, the purely mechanical application of (R1) and (R2) will determine what may be termed the *abductive success* of that explanation: an explanation is abductively successful if and only if the explanandum does indeed deductively follow from the explanans. Decisions about abductive success are therefore routinized to the extent that deductive logic can be routinely and unproblematically applied. However, once abductive success has been established, questions concerning the worth, value, usefulness, or appropriateness of the proposed explanation arise. After the determination of abductive success, only the explanandum and the logical relation between explanans and explanandum need to be known. On the basis of this information alone, i.e. under consideration only of (R1) and (R2) the latter sorts of judgement cannot be made. However, if our considerations include reference to some sort of background knowledge (maybe in the more particular form of demands for information), then the scope of the explanandum, C, can inspire what one might call *abductive awe*: the very fact that abductive success has been achieved, that someone was able to contrive an explanans from which C would follow as a matter of course, becomes a prominent creative and intellectual achievement[22]. Often, immediate enthusiasm for a new theory can be attributed to abductive awe. A remark by physicist Georg Christoph Lichtenberg concerning LeSage's atomistic hypothesis as to the cause of gravitational force is a splendid exemplification of abductive awe:

> If it is a dream, then it is the grandest and most sublime dream ever dreamt and one with which we can fill a gap in our books, a gap that can only be filled by a dream. And if only the dream is coherent and one does not deviate from the rules of correct analogy, then it could be the truth itself or stand in for it.[23]

Newton, of course, had coined his notorious '*hypotheses non fingo*' especially with respect to hypotheses like LeSage's. Accordingly and all his awe notwithstanding, Lichtenberg referred to that explanation as a dream rather than a scientific hypothesis. Both, Newton's dictum and Lichtenberg's label put into question the scientific status of the attempt at explaining gravitation. Thus, regardless of abductive awe, something like (R3) has to be included in our considerations in order to arrive at a preliminary evaluation of the proposed explanation as a scientific explanation:

> (R3) The explanans must have empirical content; i.e., it must be capable, at least in principle, of test by experiment or observation.

Abduction marks the end of the process of *invention* or *discovery*: an explanatory hypothesis has been proposed. Concerning the scientific status of the explanatory hypothesis, we subsequently ask whether scientists should (or should be permitted to) *entertain* the hypothesis. Whether a proposed explanation is abductively successful or not can be determined by applying the simple test procedure provided by (R1) and (R2). In contrast, whether a scientist should entertain a hypothesis is a normative question. As such, it gives rise to two component issues:

(A) can 'scientific' be characterized by a single (complete) set of conditions, or can one distinguish different kinds of entertainability (in terms of plausibility, predictive power etc.), parallel e.g. to Lakatos's distinction of acceptabilities?[24]

(B) can one arrive at this complete or incomplete set of conditions by a priori reasoning or is it only revealed by scientific practice in any given instance or any series of given instances?

Any way of answering these questions falls under the jurisdiction of principle (PI) that was adopted in the previous chapter (see above, p. 30). A direct application of that principle leads to the following tentative verdict on (A) and (B):

(A & B:) The philosophical analysis which leads to a minimally restrictive and incomplete set of (normative) conditions for 'empirical content' will be supplemented in the course of scientific practice by decisions which may, in effect, establish various modes of entertaining a hypothesis[25].

Before we turn to the various ways in which scientists may thus *supplement* the insufficiently restrictive conditions on 'empirical content', those *minimal constraints* have to be formulated. Indeed, there are only so many ways of going about the determination of 'empirical content':

(i) A proposition has empirical content (or is empirically significant) if its truth and the truth of its negation are logically possible (and if its truth is physically possible);

(ii) a proposition is empirically significant if there are (empirical) tests of that proposition.[26]

For the purpose of designating a subset of '*scientific* explanations' among all 'explanations', (i) is clearly insufficient, if not nearly vacuous: if only logical possibility is considered, (i) merely equates contingency with empirical significance. But even if physical possibility is included as a further demand, (i) rewards and encourages the formulation of general laws which in their antecedents make no reference to anything in the domain of accepted physical truths, i.e. formulations such as "Whenever God is angry, an earthquake will occur". It is due to this blatant insufficiency of (i) that only (ii) should be considered, which, of course, is just what Hempel and Oppenheim had in mind. (ii), however, raises further questions, firstly and most conspicuously:

(C) what is a 'test'?

In accordance with (A & B:) we should expect that there may be various complete answers to (C) which share a common core. An intuitive expression of that common core might state: a test is some sort of operation performed on a test-statement which has been deduced from premises in such a way that the general law of the explanans is an indispensable part of the premises, and such that, as a result of the operation, a truth-value is assigned to the test-statement[27]. The problem of what is an empirically significant proposition thus shifts to the question: what is a test-statement[28]? Accordingly, our answer to (C) will have to include an answer to this latter question. Now, insofar as the test-statement is arrived at by deduction from the explanatory hypothesis, the explication of (R3) corresponds to 'deduction' as the second stage in Peirce's theory of inquiry[29] in such a way that the specification of 'test-statement' will be tantamount to the formulation of conditions on the outcome of 'deduction'. But according to Peirce, regardless of its outcome, 'deduction' can be a lengthy, complicated, creative, and intuitive process, involving imprecise observations on a model of a hypothetical state of things:

> In deduction, or necessary reasoning, we set out from a hypothetical state of things which we define in certain abstracted respects. Among the characters to which we pay no attention in this mode of argument is whether or not the hypothesis of our premises conforms more or less to the state of things in the outward world. We consider this hypothetical state of things and are led to conclude that, however it may be with the universe in other respects, wherever and whenever the hypothesis may be realized, something else not explicitly supposed in that hypothesis will be true invariably. [...] All necessary reasoning without exception is diagrammatic. That is, we form an icon of our hypothetical state of things[30] and proceed to observe it. This observation leads us to suspect that something is true, which we may or may not be able to formulate with precision, and we proceed to inquire whether it is true or not. For this purpose it is necessary to form a plan of investigation and this is the most difficult part of the whole operation.[31]

Given Peirce's construal of 'abduction', his view of deduction as such an intricate process makes a tacit presupposition: by definition, an explanatory proposition resulting from abduction deductively entails that if certain initial conditions hold, then the explanandum will hold; Peirce's view of deduction thus considers irrelevant and excludes this trivially entailed deductive consequence and shoots for something *else* that can be deduced with the help of that explanatory proposition. But regardless of whether we want to adopt this additional constraint on what counts as an outcome of deduction, precise control over the success of the operation becomes possible - as with abduction - only at its very end. However, if Peirce is right, it may take much time and effort until a significant test-statement has actually *been produced.* We are thus faced with the more general issue, to what extent the empirical significance of a proposition should depend on the possibly time- and energy-consuming vagaries of deduction.

> (D) For a proposition to be empirically significant: what relation has to obtain between it and a test-statement that is deduced from it?

Once questions (C) and (D) are answered, presumably certifying the entertainability of a given explanatory hypothesis, we move to the last stage in the development of a theory. After the invention of a hypothesis by abduction and the decision to entertain it in the course of deduction, the final question of its acceptance by induction arises.

> Having, then, by means of deduction, drawn from a hypothesis predictions as to what the results of experiments will be, we proceed to test the hypothesis by making the experiments and comparing those predictions with the actual results of the

> experiment. [...] When, however, we find that prediction after prediction [...] is verified by experiment [...] we begin to accord to the hypothesis a standing among scientific results. This sort of inference it is, from experiments testing predictions based on a hypothesis, that is alone properly entitled to be called *induction*.[32]

The truth of the sentences is determined in respect to the outcomes of tests. If there should be various notions of 'test' (and various entertainabilities), the inductive operations performed on the outcome of tests are in relation to various kinds of acceptance, e.g. acceptance as true, acceptance as a strongest potential explanation, etc.[33] . Thus, the problems of 'induction' are largely derivative, reflecting the issues surrounding 'deduction'. The one and only new issue introduced by induction is the problem traditionally associated with it:

> (E) what sort of intellectual operation(s) on the outcomes of tests should be used to warrant a decision on acceptance or rejection of the explanatory hypothesis?

This formulation of (E) also leaves room for alternative ways of proceeding. In order to arrive at a theory of alternatives in science, we will have to consider (C), (D), and (E), hopefully in such a way that their shared focus becomes clear enough for us to actually speak of *one* theory of alternatives in science.

ii.

It is little appreciated that in *Testability and Meaning* Rudolf Carnap presented a theory of unilinear (non-revolutionary) development of scientific knowledge and with it the beginnings for a theory of alternatives in science[34]. Indeed, these aspects of the article are relatively indiscernible given that its rather more overt purpose was to formulate minimal conditions on 'empirical significance'. In effect, responding to questions (C) and (D), it maintained

> that all scientific terms could be introduced as disposition terms on the basis of observation terms either by explicit definitions or by so-called reduction sentences, which constitute a kind of conditional definition.[35]

This statement of purpose encapsulates the main features of the article: the emphasis on terms rather than propositions, the distinction between observational and 'scientific' terms, and the distinction between explicit definition and reduction sentences which are "a kind of conditional definition".

Though fairly inconspicuous, the shift of emphasis from propositions to terms provides an effective device in the explication of (R3) and in response to (C) and (D): after discussion of the empirical significance of terms has provided us with an answer to (C) ("What is a test or what is a test-statement?"), the shift back from terms to propositions coincides exactly with the ensueing discussion concerning the empirical significance of propositions and the answer to (D) ("For a proposition to be empirically significant: what relation has to obtain between it and a test-statement that is deduced from it?").

Thus, we start off by looking at the language of science as a collection of terms which can be taken one by one. The problem of empirical significance arises whenever a newly proposed explanans contains terms which are not yet part of that language.

> The introduction of a new term into a language is, strictly speaking, the construction of a new language on the basis of the original one.[36]

Here, a language which contains only observation terms is taken as the original language of science: it is empirically significant throughout and all further terms are introduced on its basis. While Carnap does discuss how elaborate this basic observational language might be or might have been[37], he stresses that at any given time the boundary between observational and 'scientific' terms is largely conventional, reflecting the present state of knowledge.

> A predicate 'P' of a language L is called *observable* for an organism (e.g. a person) N, if, for suitable arguments, e.g. 'b', N is able under

> suitable circumstances to come to a decision with the help of few observations about a full sentence, say 'P(b)', i.e. to a confirmation of either 'P(b)' or '-P(b)' of such a high degree that he will either accept or reject 'P(b)'.[38]

This explication of 'observable' draws "an arbitrary line between observable and non-observable predicates in a field of continuous degrees of observability"[39]. Thus, we can now draw the line between a basic language and a new language to be constructed upon it: the basic language consists of observational terms and those 'scientific' terms which have already been constructed upon the observational basis, while the new language is the basic language plus one additional 'scientific' term.

The decisive next step consists in elucidating the notion of 'construction of one language on the basis of another'. It is here, where that third main feature of Carnap's article, the distinction between explicit definition and physical reduction comes into play.

> If we wish to construct a language for science we have to take some descriptive (i.e. non-logical) terms as primitive terms. Further terms may then be introduced not only by explicit definitions but also by other reduction sentences. The possibility of *introduction by laws*, i.e. by physical reduction, is, as we shall see, very important for science, but so far not sufficiently noticed in the logical analysis of science. On the other hand the terms introduced in this way have the disadvantage that in general it is not possible to eliminate them, i.e. to translate a sentence containing such a term into a sentence containing previous terms only.[40]

Carnap distinguishes two way of reducing terms of the new language to the terms of the old one: definition and physical reduction. Reduction by definition allows for the elimination of the new terms; defined terms do not create the power to express information which is not already expressible in the basic language. Terms introduced by law cannot be so eliminated and do come with the power to express new information. We shall see on the other hand that an explicit definition fully specifies the meaning of a term while physical reduction provides only a conditional definition, leaving the meaning partly undetermined pending further inquiry. Definition can then be construed as a limiting case of physical reduction: when the conditions upon which a conditional definition rests are trivially or contingently always fulfilled, physical reduction becomes definition.

For a step by step review of this juxtaposition, we should first see why it is that explicit definition cannot be the only mode of reduction in science. A definition of a property D has the following form:

$$D_x \Leftrightarrow \ldots.x\ldots.$$

Here, D is the definiendum and the sentential function '....x....' the definiens[41]. Now consider a case where the definiendum is a conditional. Such might be the case when we want to define a dispositional property like 'soluble' (i.e. 'soluble in water'). A substance can be said to have this property if and only if it dissolves when immersed in water. Indeed, most if not all presumed structural properties of nature are designated by conditional statement[42] so that therefore, if we choose to name that property, the term chosen for that name will be defined by that conditional[43]. But let us return to the simple case of 'soluble': a substance is soluble if and only if it displays response R ('dissolves') when subjected to stimulus S ('is immersed in water').

$$D_x \Leftrightarrow (S_x \Rightarrow R_x)$$ [44]

Conceivably, as a first complaint about this definition we might assert that it does not at all give us the meaning of 'soluble': it does not tell us either what 'soluble' *is* or what we *mean* as we use that word in different contexts; instead it only tells us when to call a substance 'soluble'[45]. This criticism, however, need not concern us here: after all, we are not searching for a descriptively accurate representation of the meaning of a term. Instead, we only need a conventional way of representing terms as empirically significant within the context of 'scientific explanation' – for which, in turn, the H-D model provides a conventional representational device. For that limited purpose, stating when a substance can be called *soluble* tells us enough about the meaning and use of that term. And of course, we shall not expect that every occurence of 'soluble' in scientific discourse can be construed as satisfying this operational definition: uses of 'soluble' outside of explanatory contexts will have to be accounted for in historical terms. Indeed, terms have histories and we should suspect that a term is used differently within the context of explanation than it is before it ever appears in a general law; and that after having been used for explanatory purposes on different occasions, its full meaning may yet again go beyond a plain operational definition[46]. Thus, criticism along the lines of meaning-analysis does not challenge the validity of reduction by definition as a representational device.

However, we still have to ask whether reduction by definition yields a *good* representational device. Remember, for instance, that the H-O model was not intended as a descriptively accurate representation of belief-states of persons who propose explanations. And yet, it was argued that we can plausibly assume that every use of 'because' commits a person to some general law such that conditions (R1) and (R2) are satisfied. Now, can we assume with equal plausibility that general laws in an explanans employ terms which are introduced by definitional reduction? This amounts to asking whether the operational definition provides a *good rule* for the attribution of property D to any given instance of *x*. Using it, we will attribute D to any piece of sugar, regardless of whether it is (or has been) immersed in water. For, if S obtains R will also obtain and D should be attributed, and if -S obtains the definiens is vacuously true and D should be attributed. Likewise, we will attribute -D to a piece of wood which has been immersed in water and failed to dissolve. So far, so good. But what about a piece of wood which was burned before ever having been immersed in water? -S obtains in its case and the definiens is vacuously true, i.e. the rule is satisfied only if we attribute the disposition 'soluble' to that piece of wood[47]. Clearly, this is a most undesirable result - and it obtains for all terms introduced in conditional form (i.e. for any dispositional property)[48]. In order to avoid this result, Carnap proposed a better rule for the attribution of property D. He replaces unconditional definition by conditional definition or physical reduction. As a result, our rule for the attribution of 'soluble' takes on the form of a bilateral reduction sentence[49]:

$$(x)(t)\ (S_{xt} \Rightarrow [D_x \Leftrightarrow R_{xt}])$$ [50]

This sentence reads "if *x* is immersed in water at *t*, *x* is soluble if and only if it dissolves at *t*". Again, the issue is not whether the sentence renders a descriptively accurate representation for the meaning of 'soluble'; but rather again, whether adopting this version of an operational definition in conjunction with the H-O model yields an appropriate representational device for 'scientific explanation' in contradistinction to mere 'explanation'. Like the definition, the bilateral reduction sentence is a rule guiding the attribution of D. Here, D will be attributed if S and R, -D if S and -R are satisfied. For those cases where -S holds and either R or -R, D cannot be determined: the reduction sentence will be satisfied regardless of whether we attribute D or -D since it is vacuously true if -S. Thus, the rule functions only under the condition that S obtains; therefore it provides a conditional definition, telling us for all those (and only those) substances which satisfy S whether properties D or -D shall be attributed to them. The conditional definition *tests for* that property, and consequently S can be termed the 'test-condition' of D. Once it is satisfied, R functions as 'truth-condition', since it decides the attribution of D or -D respectively[51]. The bilateral reduction tests for property D, where 'test' can be thus defined:

$$\text{Test}_{Dx} =_{Df} (S_x \mathbin{\&} [R_x \vee -R_x])$$

That is, the conjunction of S and either R or -R constitutes a test for D. And yet, the bilateral reduction sentence is not an empirical proposition, since it rules out only one state of affairs, namely the conjunction $(S \& R \& S \& -R)$[52], i.e. it prescribes only what is a logical truth in the first place, namely $(R \vee -R)$. While it *tests for* a property, the test as just defined is not a *test of* the reduction sentence.

However, the *test for* a property can function as a *test of* a proposition once predicate D is *used* referentially in respect to some specified domain. For example, the proposition $(x)\ (Sg_x \Rightarrow D_x)$ ("if *x* is sugar, then it is soluble") is tested by bringing about facts which satisfy the conjunction of S and either R or -R. And for wood we will find that the proposition $(W_a \Rightarrow D_a)$ is tested only if W_a is a piece of wood which has been immersed in water, while $(x)\ (W_x \Rightarrow D_x)$ is falsified by modus tollens as soon as we discover at least once the conjunction of W_a, S_a and $-R_a$. Thus, the definition of '$test_{Dx}$' can serve as a preliminary answer to our question (C) ('what is a test?'), presupposing that there is a bilateral reduction sentence which determines (tests for) D or -D *and* a proposition in which D or -D is *used*, i.e. a proposition *tested by* the determination of D or -D. And thus, we can give a somewhat more complete characterization of 'test'[53]:

(C:) $Test_{Dx} =_{Df} (S_a \& [R_a \vee -R_a])$
I.e: A proposition which predicates property D to a substance undergoes a test if and only if the

following conditions are fulfilled:

(1) there is a bilateral reduction sentence for the predicate, i.e a procedure which tests for D or -D such that if (S & R) then D, and if (S & -R) then -D and such that D or -D cannot be attributed if -S;

(2) the test-conditions stated in the stimulus sentence are observed to obtain (or have been realized);

(3) it is determined that either R or -R obtains;

(4) aside from logical connectives, the terms appearing in the stimulus- and response-sentences are observable or have been introduced on the basis of observation terms.

The test has 'falsified' the proposition if, by virtue of the reduction sentence, it either determines D where the proposition predicated -D or vice versa[54]. If the test has not falsified the proposition, it has 'confirmed' it. If the facts corresponding to S are not merely observed to hold but have been actively *realized* by the persons conducting the test, the test can be called an 'experiment'.

If we now link this notion of 'test' to our discussion so far of what constitutes 'scientific activity', we find that conditions (R1) and (R2) (which formulate the hard core of the deductive nomological model of explanation) designate a privileged context for terms, i.e. the context of explanation and prediction, of general laws and their instantiations[55]. In order to restrict this context to *scientific* explanation and prediction, i.e. to establish the empirical significance of those laws, we have now added the further minimal requirement that all non-logical terms appearing in them shall have been introduced by bilateral reduction sentences on the basis of the observation language, or rather, that those terms shall be representable as having been so introduced. This additional requirement establishes or makes explicit the test-conditions for the general law.

Yet it is still far from clear what constitutes the *legitimate use* of a term introduced by a bilateral reduction sentence, i.e. the relationship between the statement of test-conditions and the empirical significance of a general law. Surely, the substitution of 'definition' by 'bilateral reduction sentence' solved the problem posed by the piece of wood which has never been immersed in water and which - by definition - would have to be called 'soluble'. In its place, however, we now have another problem: Can we call a piece of sugar 'soluble' which has never been immersed in wàter?

Consider the general law (x) $(Sg_x \Rightarrow D_x)$ which reads "All sugar is soluble". This law predicates 'soluble' of a great many things which have never been immersed in water, some of which never will be (since they no longer exist). All these substances have not been tested for D or -D. In respect to these substances, the law is neither confirmed nor falsified. Therefore, our considerations concerning the introduction of empirically significant terms have to be supplemented by a series of simple decisions concerning their use[56]. These decisions have to be guided by the principle (well-known by now) that scientific activity should not be overly restricted, i.e. that liberal use of terms that have been introduced by bilateral reduction sentences should be allowed for in agreement with scientific practice. Indeed, Carnap has tacitly followed this principle as we can see from the way in which he argues for the decisions that he recommends on the use of empirically significant terms.

First of all, we should remember that we are dealing here with minimal constraints on entertainablity and not on acceptance. The problem before us does not concern inductive rules[57]. Therefore, we have all along been able to make some tacit assumptions on the relation of conditions for meaningfulness and use. For instance, 'observation language' was defined as a language consisting mostly of *observable* rather than e.g. manifestly observed predicates. And similarly, we require that to be empirically significant a proposition has to be *testable* rather than *tested*, i.e. merely that a test *can* be performed[58]. The decision to consider testability instead of testedness does not by itself solve our problem concerning the general law which attributes 'solubility' to everything consisting of sugar. After all, many sugary things do not exist anymore and are therefore not even 'testable'. But to require that *every* substance to which a property has been attributed shall be testable in respect to that property, i.e. to require complete testability would amount to outruling all general laws, i.e. all propositions containing the affirmation of a universal quantifier or the negation of an existential quantifier. This result does not agree with scientific practice[59]. Our second decision, therefore, is to require only incomplete testability, i.e. to require only that at least one instantiation of the law is testable. And our third and final decision pertains to a hypothetical state of the following sort: while we can test for solubility in water by simply taking a substance and immersing it in water, consider the case where we want to attribute to substances solubility in fluid *Q* which is known to exist only on another, as of yet inaccessible planet? In the case of 'solubility in water', we can realize test-condition S. While in the case of 'solubility in *Q*' we could determine by observation whether S or -S and R or -R are satisfied, we cannot now (experimentally) realize them. In light of this, should we require that 'testability' shall be construed broadly in terms of 'observability' or narrowly in terms of 'realizability'? Again, for the sake of non-restrictiveness and in view of scientific practice, we opt for the broad construal[60].

This series of decisions introduces a very liberal criterion for the empirical significance of propositions that are proposed by abduction as candidates for 'scientific explanation'. They also constitute an answer to (D), i.e. to the question: "For a proposition to be empirically significant: what relation has to obtain between it and a test-statement that is deduced from it?"

> (D:) If predicate D is any non-logical term that is not part of the observation language but which is introduced by a bilateral reduction sentence, then: a proposition which uses D is empirically significant if and only if it is incompletely testable, i.e. if a test but not necessarily an experiment [as defined in (C:)] can be performed in respect to D for at least one of its instantiations as prescribed by the proposition.
> I.e: the proposition is empirically significant once one attribution of D or -D is satisfiable: as soon as for one instance the observability of the test-conditions can be assumed.[61]

iii.

(D:) provides a fairly general minimal specification of (R3). As such, it seems almost uncontestable. As we shall now see, those minimal conditions for the entertainability, i.e. empirical significance of explanatory hypotheses are far from sufficient. Taking them as such would establish a very liberal criterion indeed. And yet, I suggest that a more stringent criterion cannot be defended, not at least on the basis of a consensus within contemporary philosophy of science. Such lack of consensus is, of course, purely contingent and may express nothing but the lack of inventiveness, interpretive skill, or willingness to be persuaded on the part of philosophers of science[62]. If I choose to consider it indicative of the existence of a *variety* of available ways of complementing the minimal specification of (R3) in the course of scientific practice, I am still prepared to admit that, after all, scientific discourse could turn out to be more homogeneous than anticipated: the suggestion that the scientific character of explanatory hypotheses is negotiated within differing modes of scientific discourse is a self-correcting empirical hypothesis about the conduct of scientists. Thus, we are on a fairly safe course and do not have to consider the insufficiency of (C:) and (D:) an unwelcome result. Instead, we should go ahead and investigate some of the ways in which the all too liberal specification of (R3) is amended case by case. To that end, we will have to see first of all on what grounds the criterion as stated so far proves to be insufficient[63]. Three classical arguments to that effect will be presented:

- It has been argued that (C:) and (D:) provide no basis for excluding a variety of deeply flawed and obviously non-scientific explanations, i.e. that a subset of 'scientific explanations' from the set of all 'explanations' is not successfully formed with the help of (C:) and (D:). On occasion, it has been concluded from this that such a subset cannot properly include any dispositional explanations at all.
- Another argument maintains that bilateral reduction

sentences cannot accomodate of transient dispositional properties, i.e. that using this method of reduction amounts to the presumption that all properties are immutable, which is thought to be an untenable assumption.

- Finally, theoretical terms have been contrasted with (pure) disposition terms. Bilateral reduction sentences treat all predicates as (pure) disposition terms and do not capture the function and meaning of theoretical terms in science and scientific explanation.

As we turn to these criticisms one by one, we should be aware that none of them invalidates (C:) and (D:) as *minimal* requirements on the empirical significance of predicates occuring in the general law of a scientific explanation, i.e. on the entertainability of explanatory hypotheses.

- One cannot infer from the insufficiency of (C:) and (D:) that they should be abandoned rather than supplemented.
- If there is a difference between immutable and transient properties such that the latter are not representable by bilateral reduction sentences, that difference may be contingent upon the state of physical inquiry itself. And it may be that on an appropriate construal of 'immutable' (vs. 'transient') only

immutable properties are indeed admissable into the context of explanation. For that context, therefore, (C:) and (D:) would still be appropriate minimal requirements.

- If theoretical terms cannot be represented by single bilateral reduction sentences, it is because the meaning and use of these predicates is far more complex than the meaning and use of (pure) disposition terms. This does not entail, however, that within the context of any given explanation a theoretical term cannot be *represented as* a pure disposition term.

Thus, only the first of these three criticisms strictly demonstrates the insufficiency of (C:) and (D:) in respect to the task of selecting 'scientific explanation' as a subset of 'explanation'. The latter two maintain on semantic grounds the inadequacy of (C:) and (D:) when it comes to fully describing scientific activity. They assert that consideration of the use and meaning of scientific terms should not be restricted to the narrow context of explanation and prediction. Instead, one should deal with the language of science and the structure of scientific theory as a whole. These two criticisms are thus at odds with our deliberate insistence so far on the narrow context of explanation and prediction which was thought to lie at the core of scientific activity[64]. It enters only now, as the limits of our self-imposed restriction become apparent: we are motivated to take special notice of, e.g., theoretical terms, as opposed to (pure) disposition terms, insofar as we seek access to the way in which scientists negotiate ways of supplementing the insufficient conditions (C:) and (D:). Indeed, the very genesis of theoretical terms may be attributable to that process of negotiation. Thus, the relevance of the latter two criticisms emerges as we extend our range of vision and start focussing on the ways in which scientists supplement (C:) and (D:).

The second criticism (concerning immutable and transient disposition terms) affords us a short and narrowly conceived illustration of this attempt to turn what was designed as a criticism of Carnap's proposals into an affirmation of (C:) and (D:) as minimal requirements on scientific explanation, and into a heuristic for the discovery of considerations that may be used to supplement (C:) and (D:). That illustration will prepare us for the more far-reaching project of scrutinizing the remaining two criticisms and finally formulating the sought-after theory of alternatives in science.

Pap and Mellor[65] have pointed to Carnap's tacit and apparently untenable assumption that dispositional properties are immutable, i.e. non-transient. Indeed, consider testing a piece of iron for the property of having positive electrical charge. Consider further that the test establishes that the conditions stated by the S- and R-sentences are satisfied for the piece of iron at the time of the test. And still further, that the disposition term is introduced by our now familiar reduction sentence:

$$(x)(t)\ [S_{xt} \Rightarrow (D_x \Leftrightarrow R_{xt})].$$

Since D_x has no time-index, we attribute that dispositional property to the piece of iron as an immutable property. The test has, once and for all, determined the piece to have this property. Yet, electrical charge is a transient property - a piece of iron can carry such charge off and on throughout its existence. Thus, we clearly have a problem here. But it is a problem that lies beyond the scope of philosophical inquiry and as such a desirable problem. For there is a very real physical difference between permanent and transient dispositional properties. That difference has to be investigated by physical scientists and should be representable rather than glossed over by reduction sentences for disposition terms. Instead of trying to find a construal of such sentences that covers both cases, we should admit deviant variations on the bilateral reduction sentence as warranted by a given state of physical knowledge. For instance, a slapdash way of fixing reduction sentences for transient dispositions consists simply in adding another time index:

$$(x)(t)\ [S_{xt} \Rightarrow (D_{xt} \Leftrightarrow R_{xt})]$$

Now, a substance has a given dispositional property just at those times at which it responds to a stimulus S so that it satisfies R. Making the attribution of a property dependent on the very act involved in its observation, seems to be a highly unsatisfactory result. Yet, this unsatsifactory way of introducing a predicate may correspond at times to an equally unsatisfactory state of physical knowledge of the subject. Consider, for instance, the case in which during routine experiments, a certain anomalous property is observed at irregular intervals. Though scientists recognize the property and know the general conditions under which it may obtain, it is irreproducible (they cannot produce it at will) and they do not know whether it is transient or permanent. It is most appropriately introduced by the unsatsifactory reduction sentence that was just stated, where R is a general description of the experimental condition and S states the anomalous observations. Clearly, at this time, one cannot use D_{xt} as an explanatory predicate; indeed (C:) and (D:) forbid its use as such. Now, in order to gain an understanding of the property, scientists may try to isolate the relevant parts of the general experimental set-up and some further initial conditions that in the experiment had not been held constant - hoping that they come upon a specific construction of the state in which the anomalous observations will unfailingly occur. Once they have found that construction, they can re-introduce the disposition predicate: in virtue of an inductively confirmed generalization ('unfailingly'), the additional time-index can now be dropped[66]. In the course of the investigation, the stimulus conditions have probably undergone significant modification, and it is quite possible that the property will now be attributed not to a substance but rather to a substance in a certain state which may have a clearly delimited extension in time. In respect to that state and those stimulus conditions, the transient dispositional property has become permanent[67]. And indeed, this result agrees quite well with our intuitive notion of explanation: we explain by pointing out that (everything else being equal) it is an immutable property of the world that the general law is true and that therefore, if the initial conditions are satisfied, immutably the explanandum should be expected to occur[68]. The introduction of disposition terms may thus have a history which becomes intelligible precisely because frictions with the standard method of introduction by bilateral reduction sentences do occur as we juxtapose, e.g., pretheoretical contexts with the context of explanation and prediction[69]. In this case, the considerations supplementing conditions (C:) and (D:) may be sought in the decision-making processes on the transition from one context to the other.

The example shows how the apparent initial insufficiency of bilateral reduction sentences is placed into historical perspective by positing the form in which they introduce predicates as a first and tentative goal of physical inquiry itself: no amount of philosophical reasoning can determine whether a given property is permanent or transient or under which conditions an apparently transient disposition can be construed as a surefire disposition[70]. And, equally, philosophical inquiry cannot *alone* determine in any given instance the question of empirical significance, i.e. whether a given explanation is a properly 'scientific' explanation. This, finally, is highlighted by the first and most compelling criticism of conditions (C:) and (D:), i.e. the demonstration that they do not successfully delimit a subset of 'scientific explanations'. This criticism is most forcefully elaborated by those who quite misleadingly deny that disposition terms can be used for explanatory purposes at all, and who therefore uphold that Carnap's method of introducing "all scientific terms [...] as disposition terms on the basis of observation terms" cannot be upheld[71].

For an elucidation of that claim, consider the following two explanations:

(I) "*a* dissolves in water because it is soluble":

general law	$(x)\ Sg_x \Rightarrow D_x$	
initial condition	Sg_a	
initial condition	S_a	
explanandum	R_a	

[Sg: 'x is sugar'; D: 'x is soluble'; S: 'x is immersed in water'; R: 'x dissolves']

This is an elliptical formulation, a complete statement requires the inclusion of the reduction sentence for D:

general law	$(x)\ (Sg_x \Rightarrow D_x)$
red. sent.	$(x)\ (S_x \Rightarrow [D_x \Leftrightarrow R_x])$
initial condition	Sg_a
initial condition	S_a
explanandum	R_a[72]

(II) "a human body which contains substance *a* falls asleep because *a* possesses dormitive power":

general law	$(x)\ (O_x \Rightarrow D_x)$
red. sent.	$(x)\ [S_x \Rightarrow (D_x \Leftrightarrow R_x)]$
initial condition	O_a
initial condition	S_a
explanandum	R_a

[O: 'x is opium'; D: 'x has dormitive power'; S:'x is introduced into a human body'; R: 'x's host body falls asleep'][73]

Both (I) and (II) meet conditions (R1) and (R2) and are thus 'explanations'. Moreover, both meet our minimal requirements for empirical significance as stated in (D:), i.e. they are both testable[74]. As far as we can tell, (I) and (II) may well be properly 'scientific' explanations, at least they have 'prima facie entertainability'. Yet, (II) is notorious for being a paradigm example of an illegitimate, non-scientific explanation. It was carefully culled from Moliere's *Malade Imaginaire* in order to show that explanations in terms of dispositional properties are preposterously ridiculous - and by implication, to show that theories which (like Carnap's) do not differentiate between good and bad, dispositional and non-dispositional explanations are deeply flawed.

> First Doctor: [...]
> Demandabo causam et rationem quare
> Opium facit dormire.
> [*He sits down.* Beralde *signals to* Argan, *the* BACHELIERUS. *He rises.* Beralde *prompts him.*]
> Argan: Mihi a docto Doctore
> Demandatur causam et rationem quare
> Opium facit dormire.
> And to that respondeo
> Quia est in eo
> Virtus dormitiva
> Cujus est natura
> Sensus tranquillizare.
> Chorus: Bene, bene, bene, bene respondere!
> Dignus, dignus est entrare
> In nostro docto corpore.[75]

The claim that dispositions cannot serve explanatory purposes presupposes that *this* dispositional explanation shares a common defect with all explanations in reference to dispositions. However, that presupposition can be challenged simply by highlighting significant differences between the 'solubility' and 'dormitive power' explanations. By the same token, we will discover clues as to how to supplement (C:) and (D:) in such a way that (I) but not (II) belongs to the subset of 'scientific explanations'. As a result we may consider the explanation parodied by Moliere as fundamentally defective, while (I) is only somewhat uninformative.

To see what is wrong with the 'dormitive power' explanation it is really not enough to consider only the formulation given in (II). Both (I) and (II) seem uninformative since they assert only that some substances display a certain response to a given stimulus because they are disposed to respond in such manner to that stimulus. I can see nothing intrinsically wrong with such exercises in apparent futility[76]. Yet, this sober assessment of (II) is not shared by Moliere's protagonists: they take a childish delight in their explanation, clouding it, as they do, in an aura of profundity. And that is precisely what is wrong with their explanation: by taking it as the discovery of some essential force, they effectively preclude further inquiry into the causal relations which obtain here. They adopt a 'mystery-raising' rather than 'problem-raising' attitude towards their dispositional explanation[77], and what should be the beginning of inquiry into, e.g., the shared molecular properties of all substances with 'dormitive power', is taken as the ultimate revelation of a principle residing within the substance before them[78].

The context for the 'solubility' explanation is quite different. Indeed, while 'soluble' is one of the favored examples when it comes to discussing disposition terms, our current state of chemical knowledge has far transcended explanations of a particular dissolution in terms of 'solubility'. A plausible, though not necessarily chemically or historically accurate scenario shows how scientists might have arrived at a far more thoroughgoing explanation of R_a[79]. Treating the disposition predicate in the general law $(x)(Sg_x \Rightarrow D_x)$ as a problem-raising predicate, scientists scrutinized all soluble substances[80] for shared properties or for distinctive traits that would allow for fruitful sub-classifications. Thus, by comparing the behavior of salt, sugar and other substances upon being immersed in water, they discovered different kinds of solubility: crystals form in some solutions and not in others; after evaporation of the liquid, the solute sometimes remains behind and sometimes it has evaporated with the solution; as opposed to salt, sugar does not dissolve well at low temperatures; by adding a molecule of salt to water, the boiling-point is raised twice as much as when one adds a molecule of sugar; etc. By further investigating the shared properties of all substances in any group of solubles, they discovered, for instance, that electricity is conducted well by all those solutions from solutes that also dissolve well at low temperatures, and that all other solutions do not conduct electricity[81]. Thus, they might divide all solutes into 'electrolytes' vs. 'nonelectrolytes'. Then, they discover that electrolytes are "substances that can dissociate into electrically charged particles called ions, while nonelectrolytes consist of molecules that bear no net electrical charge"[82]. Thus, soluble substances consisting of molecules which bear net electrical charge have the disposition to dissociate into ions upon being immersed in water. Scientists can now presumably tell for all soluble substances whether they are electrolyte or nonelectrolyte solutes merely by looking at their molecular properties. As a result of this development in chemistry, we can now define 'soluble'

$$D_x \Leftrightarrow (E_x \vee NE_x),$$

where E stands for 'is electrolytically soluble' and NE for 'is nonelectrolytically soluble'. And for E and NE respectively, we have these first bilateral reduction sentences:

$$(x)\ (S_{1x} \Rightarrow [E_x \Leftrightarrow R_{1x}])$$

$$(x)\ (S_{1x} \Rightarrow [NE_x \Leftrightarrow -R_{1x}]),$$

where S_1 means 'x is immersed in water at a low temperature' and R_1 'x dissolves well'[83]. They were soon supplemented by

$$(x)\ (S_{2x} \Rightarrow [E_x \Leftrightarrow R_{2x}])$$

$$(x)\ (S_{2x} \Rightarrow [NE_x \Leftrightarrow -R_{2x}]),$$

where S_2 stands for 'x is dissolved in water and the solution of x is tested for its (electrical) conductivity' and R_2 for 'x conducts electricity well'. And finally, on the highest level of analysis so far, we have:

$$(x)\ (S_{3x} \Rightarrow [E_x \Leftrightarrow R_{3x}])$$

$$(x)\ (S_{3x} \Rightarrow [NE_x \Leftrightarrow -R_{3x}]),$$

where S_3 means 'x dissolves in water and the molecules of x have beforehand been subjected to a procedure determining their net electrical charge', and R_3 'before having dissolved in water the molecules of x bore net electrical charge'. And, in conclusion of our summary, we can now give a more thoroughgoing explanation of R_a ('a piece of sugar *a* dissolves in water')[84]:

(Ib) "*a* dissolves in water because it is (non-electrolytically) soluble", i.e:

general law	$(x)\ (Sg_x \Rightarrow NE_x)$
definition	$(x)\ (D_x \Leftrightarrow [E_x \vee NE_x])$
red. sent.	$(x)\ (S_x \Rightarrow [D_x \Leftrightarrow R_x])$
initial condition	Sg_a
initial condition	S_a
explanandum	R_a

[Sg: 'x is sugar'; NE: 'x is nonelectrolytically soluble'; D: 'x is soluble'; E: 'x is electrolytically soluble'; S: 'x is immersed in water'; R: 'x dissolves'][85]

Explanation (Ib) still represents only an intermediate step towards a full explanation for R_a as it could be given today[86]. But already we can make a number of observations about (Ib) which we could not make about either (I) or (II) and which set apart an explanation in reference to solubility from the explanation parodied by Moliere.

- In virtue of the various reduction sentences for E and NE, we cannot only explain (or predict) that the piece of sugar dissolves, but furthermore that the resulting solution does not conduct electricity well and that the sugar will not dissolve easily at low temperatures.
- Explanation (Ib) is no less dispositional than (I) or (II). However, it contains what one might consider higher- and lower-level dispositions, i.e. 'nonelectrolyte solubility' and 'solubility'.
- It now turns out that in respect to more elaborate (Ib) the explanation put forth in (I) was nothing but an abbreviated way of talking.

Also, it should be noted that our scientists arrived at (Ib) only because (I) was neither discarded nor elevated to the status of an ultimate revelation, i.e. because D was treated as a problem-raising predicate. And finally, we may surmise that the use of 'soluble' in (I) appears non-problematic to us largely because it is an elliptical expression, that is, because scientific understanding of the processes involved goes far beyond (I), and because we feel confident that accordingly we could give a more comprehensive statement of them. Our scenario so far can now be summarized: scientists decided to treat D as a problem-raising disposition term and subsequently investigated the problem raised by it; and as a result of their investigation D was rendered non-problematic as they arrived at higher level dispositions E and NE, towards which in turn they can now adopt either a problem- or mystery-raising attitude[67].

The difference between the two dispositional ('soluble' and 'dormitive power') explanations consisted at first only in the different attitudes adopted towards them. Due to the adoption of these attitudes, the difference grew to be more complex. As opposed to the one-dimensional, mystery-raising dormitive power explanation, our explanation of the dissolution of a piece of sugar now involves various levels of dispositionality and a non-problematic notion of 'soluble'. So, if there were criteria which might supplement conditions (C:) and (D:) so that we could unequivocally distinguish scientific from non-scientific, empirically significant from empirically insignificant explanations, the juxtaposition of (Ib) to (II) should afford us important clues. The interpretation of these clues might consequently enable us to shed further light on the difference between (I) and (II). More precisely then, to arrive at criteria that would supplement (C:) and (D:), we would have to make explicit .

(i) the structural differences between explanations of types (Ib) and (II), and
(ii) the reason that explanation (I), and not (II), was developed into full-fledged (Ib); i.e. ideally, why (I) and not (II) is disposed to evolve into an explanation of type (Ib).

As far as (i) is concerned, we shall see that there are at least two different interpretations of the difference between (Ib) and (II) and that the choice of interpretation is highly contingent upon scientific inquiry: depending on the outcomes of elaborate negotiation, the difference between the two explanations either proves (Ib) as a certifiably scientific accomplishment or remains tenuous and tentative. Due to this contingency, we also will not find any intrinsic or structural differences between (I) and (II). Both are plain dispositional explanations and only the choice to treat (I) as problem- and (II) as mystery-raising renders (I) and not (II) entertainable as a 'scientific explanation' even if its empirical significance has not been demonstrated in full.

In our step by step construction of scientific activity, we thus arrive for the first time at momentous junctures which require more than highly routinized decision-making. After all, checking for abductive success and satisfaction of conditions (C:) and (D:) is an almost mechanical procedure, requiring no more than knowledge of deductive logic and the list of terms which at any given time belong to the language of science. But, while treating (I) as a problem-raising explanation seems to be the only choice to make in the context of modern science, it nevertheless requires some *bon sens* or *Urteilskraft*[88] to make that choice and act in accordance with it. That it takes more than declared intent to make this choice becomes apparent when controversies arise as to whether, e.g. evolutionary biologists do *in effect* treat certain explanatory theories in a mystery-raising manner, regardless of their professed commitments. Scientists often have to show by argument that indeed they are treating explanatory concepts in problem-raising ways. And similarly, all the other choices on the road to the acceptance of (Ib) and of 'soluble' as a non-problematic disposition term involve scientific negotiation beyond routinized comparisons of standards and criteria on the one hand, theories on the other.

This, then, is our present standing in relation to our overall task: proposed explanatory hypotheses are the product of scientific activity, and routine decison-making ensures that scientific hypotheses are abductively successful and meet conditions (C:) and (D:). From scientific activity emerges at this point a set of alternatives concerning the further treatment of hypotheses. So far, this set consists only of the choice between treating an hypothesis as either problem- or mystery-raising. However, as we undertake the promised comparison of (Ib) and (II), we shall soon arrive at a more complete characterization of the alternatives at stake.

iv.

Assuming that explanation (I) is just like (II) in that it expresses all that we can say about the dissolution of the piece of sugar, we can juxtapose (I) and (Ib) using several overlapping formulations:

- 'Soluble' is a problem- or mystery-raising term in (I), non-problematic in (Ib).
- Explanation (I) may or may not be scientific, (Ib) surely is.
- There is only one reduction sentence for 'soluble' in (I). Explanation (II) contains a definition and a reduction sentence for 'soluble' and has at its disposal at least three reduction sentences for 'nonelectrolytically soluble' as well as 'electrolytically soluble' (the defining terms of 'soluble').
- In (I), the ascription of 'solubility' to an object serves to explain only why it dissolves. Knowing that

an object is nonelectrolytically soluble permits us to make three independent predictions about solubility, conductivity of the solution, and solubility at low temperatures.

Each of these formulations expresses that (Ib) differs from (I) insofar as 'soluble' has become well-entrenched in a theoretical language[89]. If this is a plausible evaluation, we should expect that the notion of 'entrenchment in a theoretical language' can be so clarified that each of the foregoing formulations is reducible to a statement like 'there is an (entrenchment) condition which is met by (Ib) and not by (I)'. This condition could then be useful as a further criterion for 'scientific explanation'. Indeed, the last of the four presumably roughly equivalent formulations affords us a clue as to what that entrenchment condition might consist in: it might require that any scientific explanation should have testable consequences other than its explananda. The connection with the notion of 'entrenchment' is readily apparent: By itself, an explanatory hypothesis in the form of a conditional general law entails only that an explanandum will be true if the initial conditions are fulfilled. The requirement that *other* testable consequences should be entailed by the explanatory hypothesis therefore makes sense only if we amend it, so that it reads:

> (D*) Any explanatory hypothesis in conjunction with appropriate background knowledge should entail testable consequences other than the explanandum which is entailed by the hypothesis and the satisfaction of the initial conditions alone.

'Entrenchment' is now 'entrenchment within appropriate background knowledge'. In some cases, the appropriate background knowledge may not be available yet. In the case of (Ib), it is provided by the various reduction sentences for E and NE. Thus, the presence of several definitions and/or reduction sentences for a single term is a sign of theoretical entrenchment precisely because it allows for the introduction of appropriate background knowledge and the deduction of further testable consequences. If we now simply add the stipulations that only an explanatory hypothesis which meets condition (D*) is (certifiably) 'scientific' and that predicates occuring in such a hypothesis are 'non-problematic', we have concluded our task of reducing the four formulations juxtaposing (I) and (Ib) to condition (D*). The compelling simplicity of the proposal to include condition (D*) is only slightly marred by two problems which we may choose to consider minor and residual for the time being, i.e. the problems that, firstly, the status of (I) is still unclear, and secondly, that we have maybe prematurely decided not only that 'soluble' is non-problematic, but also 'nonelectrolytically soluble'.

In effect, we have abandoned or modified one of our basic assumptions, the assumption concerning the symmetry of explanation and prediction. While we still maintain that explanation and prediction share the same logical form and that what happens to be an explanation in these circumstances could also be a test or prediction under different circumstances, with (D*), *pragmatic asymmetry of explanation and prediction* has actually become a demand to be fulfilled by 'scientific explanations'[90]. Hempel and Oppenheim already envisioned this pragmatic asymmetry as a temporal or epistemic asymmetry modeled on any isolated instantiation of the H-O model. Without going quite far enough as to include the theoretical entrenchment of terms within the body of background knowledge, they write:

> Let us note here that the same formal analysis [...] applies to scientific prediction as well as to explanation. The difference between the two is of pragmatic character. If *E* [the explanandum] is given, i.e. if we know that the phenomenon described by *E* has occured, and a suitable set of statements $C_1, C_2, \ldots, C_k$ [the initial conditions], $L_1, L_2, \ldots, L_r$ [the general laws] is provided afterwards, we speak of an explanation of the phenomenon in question. If the latter statements are given and *E* is derived prior to the occurence of the phenomenon it describes, we speak of a prediction.[91]

What Hempel and Oppenheim fail to take into account is that both our explanations (I) and (II) (the explanations in terms of solubility and dormitive power) are explanatory and predictive in this elementary sense. Moreover, as it is the crucial feature of prediction that it makes reference to

phenomena which are as of yet unknown, (D*) should be adopted for yet another reason: Hempel and Oppenheim conflate two distinct steps when they say that an explanation presuppes knowledge only of an *E* for which we subsequently discover a suitable general law and a set of initial conditions. For we are not likely to discover these two things all at once and in such manner that the initial conditions just happen to be instantiations of the antecedent of the general law. On the contrary, the general law is chosen as suitable because we have observed that certain phenomena regularly concur with the explanandum and that they might do a good job as initial conditions. Thus, the explanatory use of a general law
$(x)\ (S_x \Rightarrow R_x)$ expresses our antecedent knowledge
of (S & R). Indeed, one might suspect that only this antecedent knowledge of the regular concurrence of
$(S_a \ \& \ R_a)$ gives rise in the first place to the formulation of the general law, or even, that only the observed concurrence of $(S_a \ \& \ R_a)$ gives rise to the notion that there is something worthy of an explanation here[72]. Now, while we can use the general law in conjunction with S_a to predict R_a, we cannot very well consider this a prediction of a surprising or even unsuspected phenomenon. The predicted R_a is unknown only in the most trivial sense that it has not yet been observed: a prediction of R_a serves as a test only in

the most trivial sense of testing for the (continued) validity of a generalization, i.e. of testing the uniformity of nature[93]. Thus, if we want our explanations to also serve as predictions of unknown facts, we should look for those facts outside of the immediate explanatory scope of the general law.

Everything points to condition (D*) as a suitable supplement to (C:) and (D:). Indeed, Popper's 'independent testability', Lakatos' 'excess empirical content' and 'prediction of novel facts', Goodman's 'projectibility', Glymour's 'method of bootstrapping' (in response to the 'problem of old evidence'), as well as countless variations on these criteria all aim for more or less satisfactory specifications of (D*)[94].

And yet is seems that in his later work Carnap is rather skeptical towards the possibility of neatly collapsing the notions of entrenchment and independent testability into a criterion (D*) for the empirical significance of scientific explanations. Instead, he shifts emphasis from the issue of empirical significance as a property of *scientific* theories to scientific activity as it manifests itself in an on-going process of negotiation and decision-making *about* the significance of entrenched terms as they occur in (undoubtedly scientific) explanations. Though hardly perceivable, the roots of this shift can already be found in

Testability and Meaning, which provides a splendid account of entrenchment and theoretical progress: theories develop and science progresses as further reduction sentences and/or definitions are introduced for scientific terms.

> If a property or physical magnitude can be determined by different methods then we may state one reduction pair or one bilateral reduction sentence for each method. The intensity of an electric current can be measured for instance by measuring the heat produced in the conductor, or the deviation of a magnetic needle, or the quantity of silver separated out of a solution, or the quantity of hydrogen separated out of water etc. We may state a set of bilateral reduction sentences, one corresponding to each of these methods. [...] If we establish one reduction pair (or one bilateral reduction sentence) as valid in order to introduce a predicate 'Q_3', the meaning of 'Q_3' is not established completely, but only for the cases in which the test condition is fulfilled. [...] We may diminish this region of indeterminateness of the predicate by adding one or several more laws which contain the predicate and connect it with other terms available in our language. These further laws may have the form of reduction sentences (as in the example of the electric current) or a different form[95]. [...] Nevertheless, a region of indeterminateness remains, though a smaller one. [...] This region may then be diminished still further, step by step, by stating new laws.[96]

Ideally, then, we will arrive ultimately at a set of bilateral reduction sentences for one predicate which will completely determine the meaning. Here, a set of physical reduction sentences achieves the status of a virtual definition[97].

> A set of reduction pairs is a partial determination of meaning only and can therefore not be replaced by a definition. Only if we reach, by adding more and more reduction pairs, a stage

> in which all cases are determined, may we go over to the form of a definition.[78]

This sounds straightforward enough. But it proves upon closer scrutiny to be the most puzzling passage in Carnap's article and in need of revision (as he himself would later realize). For, what happens as we entrench a predicate by adding further reduction sentences, and as we presumably approach the point at which the set of reduction sentences can be somehow transformed into a definition?

To be sure, there is a rather clear-cut case in which the addition of further reduction sentences permits scientists "to go over to the form of a definition". In that case, however, further entrenchment consists in the establishment of a biconditional (instead of a plainly conditional) relation and has nothing to do with narrowing the margin of indeterminacy. And furthermore, if we were to conceive of (D*) along the lines provided by this special case, it would end up as far too restrictive a condition. The case is provided by physical laws which are conjunctions of conditionals which are representable in the form of an equation. Newton prepares the formulation of such a law in his 1672 paper on optics:

> To the same degree of Refrangibility ever belongs the same colour, and to the same colour ever belongs the same degree of Refrangibility. The *least Refrangible* Rays are all disposed to exhibit a *Red* colour, and contrarily those Rays, which are disposed to exhibit a *Red* colour, are all the least refrangible: So the *most refrangible*

> Rays are all disposed to exhibit a deep *Violet Colour*, and contrarily those which are apt to exhibit such a violet colour, are all the most Refrangible. And so to all the intermediate colours in a continued series belong intermediate degrees of refrangibility. And this Analogy 'twixt colours, and refrangibility, is very precise and strict; the Rays always either exactly agreeing in both, or proportionally disagreeing in both.[99]

In its simplest conceivable form, the resulting equation might look like this:

$$a = kb$$

If k is a known constant, *a* is virtually defined as *kb*, and *b* as *a/k*. At the same time, the equation entails or consists of three predictive or explanatory conditionals about the measured values V for a and b.

$$(x)\ (Va_x \Rightarrow [Vb_x = Va_x/k])$$
$$(x)\ (Vb_x \Rightarrow [Va_x = kVb_x])$$
$$(x)\ ([Va_x \ \&\ Vb_x] \Rightarrow [k = Va_x/Vb_x])$$ [100]

Presuming that the equation encapsulates Newton's discovery about refrangibility and the disposition to exhibit certain colours, it introduces a dispositional property shared by all light-rays. The property D(a=kb) disposes a ray so that its refrangibility and its disposition to exhibit color concur in such a way that there is a precise and strict relation between degree of refrangibility and the disposition to exhibit a certain color[101]. Since identity (in an equation) is extensionally equivalent to a biconditional, we can simply define D(a=kb) without having to fear the counter-intuitive result of having to attribute the defined predicate whenever the antecedent of the definiens is not satisfied.

$$D(a=kb)_x \Leftrightarrow (Va_x = kVb_x)$$

Of course, one could also introduce D(a=kb) as the conjunction of three reduction sentences, each corresponding to one of the general laws stated above.

$$(x)(Va_x \Rightarrow [D(a=kb)_x \Leftrightarrow \langle Vb_x = Va_x/k\rangle])$$
$$(x)\ (Vb_x \Rightarrow [D(a=kb)_x \Leftrightarrow \langle Va_x = kVb_x\rangle])$$
$$(x)\ ([Va_x \ \&\ Vb_x] \Rightarrow [D(a=kb)_x \Leftrightarrow \langle k = Va_x/Vb_x\rangle])$$ [102]

We can easily switch back and forth between this conjunction of reduction sentences to the definition of D(a=kb). Our ability to do so has nothing to do with the sheer number of reduction sentences and a presumed narrowing of the region of indeterminacy. Indeed, all cases of $-(Va_x \vee Vb_x)$ still belong to this region. And yet, in contrast to Carnap's assertion, no further reduction sentences are required to "go over" to the definition, since the definitional form is yielded simply by the particular (biconditional) structure of the general law. To have biconditional general laws is a rare and fortunate privilege accrued within certain areas of some of the natural sciences. Therefore, if undue restrictiveness is to be avoided, a demand for this structure cannot be written into (D*).

Thus, Carnap must have had something else in mind when he conceived of theory development in terms of a progressive accumulation of reduction sentences which would ultimately lead to a definition[103]. The notion of narrowing the region of indeterminacy of meaning provides the clue. The meaning of D, introduced by bilateral reduction sentence $(S \Rightarrow [D \Leftrightarrow R])$, is undetermined for $-S$. And a set of bilateral reduction sentences for D leaves it undetermined for

$$-(S_1 \vee S_2 \vee \ldots \vee S_n)$$

. The meaning of D is completely determined when

$$(x)\ (S_{1x} \vee S_{2x} \vee \ldots \vee S_{nx})$$

is true. At that time, the reduction sentences "may be replaced by the definition"[104]

$$D \Leftrightarrow [(S_1 \,\&\, R_1) \vee \ldots \vee (S_n \,\&\, R_n)].$$

This procedure creates the impression of an almost organic process of growth which culminates in a smooth transition from reduction sentences to definitions: the accumulation of knowledge coincides entirely with the emergence of meaning. This impression, however, is somewhat at odds with another passage in *Testability and Meaning* and very much at odds with Carnap's later development of the theme in *The Methodological Character of Theoretical Concepts*. Consider our reduction sentences for NE:

$$- \quad (x)\ (S_{1x} \Rightarrow [NE_x \Leftrightarrow -R_{1x}])$$
$$- \quad (x)\ (S_{2x} \Rightarrow [NE_x \Leftrightarrow -R_{2x}])$$
$$- \quad (x)\ (S_{3x} \Rightarrow [NE_x \Leftrightarrow -R_{3x}])$$

[NE: 'x is nonelectrolytically soluble'; S_1: 'x is dissolved in low-temperatured water'; R_1: 'x dissolves well at low temperatures'; S_2: 'x is dissolved in water and the solution of x is tested for its (electrical) conductivity'; R_2:'x conducts electricity well'; S_3: 'x dissolves in water and the molecules of x have beforehand been subjected to a procedure determining their net electrical charge'; R_3: 'the molecules of x bore net electrical charge before having dissolved in water'.]

Each of these sentences corresponds to a general law concerning all substances to which we attribute NE. In *Testability and Meaning*, Carnap makes the following remark about the cumulative introduction of such laws and their corresponding reduction sentences[105].

> This region [of indeterminacy of meaning] may then be diminished still further, step by step, by stating new laws. These laws do not have the conventional character that definitions have; rather are they discovered empirically within the region of meaning which the predicate in question received by the laws stated before. But these laws are extended by convention into a region in which the predicate had no meaning previously; in other words, we decided to use the predicate in such a way that these laws which are tested and confirmed in cases in which the predicate has a meaning, remain valid in other cases.[106]

In other words, we have to *decide* whether or not all three reduction sentences for NE actually introduce one and the same concept. For,

> strictly speaking, for one concept no more than one test procedure must be given. If we specify, say for "electric charge" three test procedures, then thereby we haven given operational definitions for three different concepts; they should be designated by three different terms, which are not logically equivalent.[107]

When we decide whether or not 'NE' designates the same property in each of the reduction sentences, we are in effect negotiating the force of the 'should' in Carnap's last sentence. There are two ways of intrepreting the sameness of a term vis-a-vis the difference, strictly speaking, of concepts that are introduced each by only one bilateral reduction sentence:

(i) We revise our reduction sentences for 'NE', acknowledging that, strictly speaking, the three sentences introduce three different concepts, as for instance
$(x)\ (S_{1x} \Rightarrow [NE_{1x} \Leftrightarrow -R_{1x}])$.
We may tentatively justify our continued use of the same term ('nonelectrolytically soluble') for NE_1, NE_2, and NE_3 by making an empirical claim stating that there is no substance to which we can attribute any of the three predicates NE_1, NE_2, or NE_3 and to which we could not also attribute the other two if we conducted the appropriate tests. And thus, we might propose three pairs of hypotheses of the following kind, which all together assert that the three concepts are co-extensive:[108]
(H$_1$) $(x)\ ([S_{1x} \ \&\ -R_{1x}] \Rightarrow [S_{2x} \Rightarrow -R_{2x}])$
$(x)\ ([S_{1x} \ \&\ -R_{1x}] \Rightarrow [S_{3x} \Rightarrow -R_{3x}])$

(ii) Following Carnap's suggestion in *Testability and Meaning*, we decide to consider the

> three reduction sentences as each contributing to a complete determination of the meaning of NE. Since, strictly speaking, each reduction sentence introduces a different concept, we adopt the form of a definition to represent the meaning of NE. Though our three sentences do not yet eliminate the region of indeterminacy for NE[109], we can commit ourselves to this preliminary formulation:
>
> $$NE \Leftrightarrow [(S_1 \& -R_1) \vee (S_2 \& -R_2) \vee (S_3 \& -R_3) \vee C].$$
>
> C is a residual, as of yet unknown disjunction of those (S & R)s which would complete the definition of NE. Thus, we do not present an empirical conjecture concerning the unified use of 'NE'. Rather, we introduce a rule for the unequivocal attribution of NE. We no longer need the bilateral reduction sentences to rule on its attribution, but we can preserve their purpose of linking that predicate to the general laws upon which they were based, by transforming them into three conditionals of the form:
>
> $(P:)\ (x)\ (NE_x \Rightarrow [S_{1x} \Rightarrow -R_{1x}])$
>
> In terms of its historical genesis, the definition of NE is based on the truth of each of these conditionals[110]. This close tie has been severed as the definition was presented: at most one of these conditionals has to be true for us to rule on the attribution of NE. Thus, if we want to preserve the information that was once represented by the *set* of bilateral reduction sentences, we have to view these conditionals as meaning postulates rather than hypotheses[111].

The difference between (i) and (ii) is best illustrated by considering the following case. Some time after arriving at the state of knowledge expressed by explanation (Ib), scientists discover a new substance *a*, dissolve it in water and find

$$(S_{1a} \& R_{1a}) \& (S_{2a} \& -R_{2a}),$$

i.e. a dissolves well at low temperatures and yet the solution does not conduct electricity well. If we have chosen option (i), we attribute '$-NE_{1a}$' and 'NE_{2a}'. At the same time, the first of our three hypotheses (H:) (on the co-extension of predicates NE_1, NE_2, and NE_3) has been falsified and we will have to look for some suitable replacement which more accurately captures the relation between the dispositions to dissolve well at low temperatures and to conduct electricity in solutions. In the course of that investigation, we may find that there is absolutely no connection whatever between the two dispositions: while conductivity is closely related, say, to the net electrical charge on molecules, solubility at low temperatures is a manifestation of, say, the dispositional property called 'diffusivity'. However, if we adopt (ii), the course of things will be quite different. The definition leads to an unequivocal attribution of NE^{112}. But the first of our meaning postulates (P:) tells us that

$$(x)\ (NE_x \Rightarrow [S_{1x} \Rightarrow -R_{1x}]),$$

which clashes with our finding of (S_{1a} & R_{1a}). Now, rather than a falsification, we have before us an anomaly: instead of investigating whether or not our definition or meaning postulates should be revised, we will explore the particular circumstances which in this case are responsible for an unanticipated observation[113]. Investigation of that anomaly leads to the finding, e.g., that some other forces are at work here which either effected solubility at low temperatures where none would be expected or diminished the electrical conductivity of the solution.

Rudolf Carnap presented the issue in very similar terms. If a predicate is introduced by a single bilateral reduction sentence which "constitutes the whole meaning of the term"[114], then

> the negative result of a test for a disposition must be taken as conclusive proof that the disposition is not present. But a scientist, when presented with the negative result of a test for a certain concept, will often still maintain that it holds, provided he has sufficient positive evidence to outbalance the one negative result. [...] The scientist will point out that the test procedure for I_o based on S and R should not be taken as absolutely reliable, but only with the tacit understanding "unless there are disturbing factors" or "provided the enviromnment is in a normal state". Generally, the explicit or implicit inclusion of such an escape clause in the description of a test procedure for a concept M in terms of a condition S and a result R shows that M is not the pure disposition D_{SR}.[115]

The pure disposition predicate which is introduced by one bilateral reduction sentence admitting no escape procedures is here juxtaposed with theoretical terms which are "never completely interpreted"[116] and for which

> the test procedure involving S and R may well

admit of exceptions in case of unusual disturbing factors.[117]

The addition of further bilateral reduction sentences and therefore its further entrenchment transform a (pure) disposition term into a theoretical term[118]: it introduces the option of pursuing either course (i) or course (ii) if a problem as the one discussed just now is encountered.

In the case of the virtual definition obtained from a law in the form of an equation, we found ourselves able to move back and forth between a set of reduction sentences and a definition as easily as one can shift from a biconditional to a pair of conditionals. In contrast, though options (i) and (ii) might each accurately represent the use of 'NE' by the scientists who proposed explanation (Ib), the example has shown them to be by no means equivalent. This is not to say that they do not on occasion and under certain conditions converge.

- (ii) converges (rather, reverts) to (i), when for whatever reason one starts contemplating a revision of the definition and decides to treat the meaning postulates as empirical hypotheses the truth or falsity of which would affect the construal of the definition;
- (i) converges to (ii) once one is convinced that hypotheses (H:) are true, i.e. once one *knows* the

predicates introduced by the different reduction sentences to be co-extensive, i.e. to designate just one property.

Identifying the condition under which (i) converges to (ii) is tantamount to specifying the condition under which we are justified in adopting (ii). Inversely, scientists who treat predicates according to (ii) therewith reveal that they have (tacitly) accepted (H:) as true. After all, a definition is supposed to be non-creative, asserting nothing that could not be asserted without it: and the transition from the collection of bilateral reduction sentences to a (preliminary or complete) definition and a collection of meaning postulates is non-creative only if hypotheses (H:) are true, i.e. if all reduction sentences are known to introduce the same concept.

Now, it is only at this point of convergence that a general law containing NE becomes independently testable and NE non-problematic. Remember that (Ib) was thought to be better than either (I) or (II) since it satisfies some version of (D*). And it does so because the general law used in the explanation of R_m ('a piece of sugar dissolves') can also be used to explain or predict e.g. solubility (or non-solubility) at low temperatures and electrical conductivity (or non-conductivity) of a solution. Thus, a general law stating that all sugar is nonelectrolytically soluble can be used to explain the dissolution of a piece of sugar and is also independently testable: we test the validity of that explanation by seeing e.g. whether sugar dissolves well at low temperatures or whether the resulting solution conducts electricity well. Each of these tests is based on one reduction sentence or meaning postulate of NE. The performance of the tests presupposes therefore that we are testing for one and the same property, i.e. that the three reduction sentences introduce co-extensive predicates. Thus, a general law is independently testable and a predicate non-problematic only at the point where (i) converges to (ii), i.e. only when the adoption of (ii) is justified or when empirical hypotheses (H:) are known to be true. But, when and whether they are known to be true is obviously not a question philosophers can decide. Scientists negotiate that issue continuously on evidential grounds: once a consensus has been reached, negotiation may be reopenend by the discovery of new evidence (as in the case of the substance which dissolved well at low temperatures and nevertheless made a nonelectrolyte solution) or by the introduction of a further reduction sentence for a predicate which will significantly expand set (H:)[119]. Moreover, negotiation concerning the truth of (H:) may not be possible on evidential grounds alone, i.e. the outcome of the negotiation may not be uniquely determined by evidence. Consider again the case presented above: those who propose

(H:) conjecturally and who pursue course (i) argue for its case in reference to the evidence provided e.g. by the very phenomenon which they perceive as a falsification of (H:); yet, those who maintain that (H:) is true and that postulates (P:) should remain untouched also argue with respect to evidence, e.g. to past experience and to evidence yet to be discovered if one follows up on their (content-increasing) ad-hoc hypotheses concerning sources of interference. Thus, negotiation of (H:) is not only in principle interminable but may also prove to be difficult and tenuous.

We can conclude our elaborate discussion of explanation (Ib) in terms of nonelctrolyte solubility by making some judgements on the fictional state of scientific knowledge which was presented here. Within the confines of our scenario, 'soluble' is now a non-problematic term. It is defined by (E v NE), and it is accompanied by the meaning-postulate that, if something is soluble as defined, then it will dissolve upon being immersed in water. NE, on the other hand, is not quite non-problematic. While the sheer number of reduction sentences proposed for NE reveals that we have preferred a problem- rather than mystery-raising attitude towards it, we cannot be sure that hypotheses (H:) are true for NE. It is likely that the general law in (Ib) is independently testable in respect to electrical conductivity or net-molecular charge on the solute molecules. But the determination of solubility at low temperatures may not be a test of the attribution of NE in the general law at all. The negotiation concerning the truth of hypotheses (H:) continues and we cannot be sure to what extent criterion (D*) has been met. We would nevertheless insist that (Ib) is far more advanced than either explanation (I) or (II) in respect to (D*), and that (Ib) is therefore a better, more significant explanation.

We have thus driven a wedge between 'entrenchment' and 'independent testability'. The notion of a 'theoretical term' is associated with 'entrenchment', while non-problematic and thus certifiably significant terms coincide with the independent testability of the laws in which these terms occur. Only the presumed truth of certain empirical laws smolders entrenchment and independent testability and makes (D*) applicable to a given law[120]. Rather than as a hard and fast criterion, (D*) serves as a guiding light or regulative ideal in scientific inquiry: it is a goal of scientific activity to propose explanations which conform to (D*). While non-satisfaction of abductive success, (C:), or (D:) effectively terminates scientific activity in respect to a proposed explanation, non-satisfaction of (D*) does not. For example, explanations proposed by 17th century chemists were abductively successful and met conditions (C:) and (D:): for instance, the key concept 'phlogiston' designated the hypothesized substance corresponding to the dispositional property of burnability, and was a well-entrenched theoretical term. Scientific negotiation accross paradigms finally led to a consensus according to which 'phlogiston' is not empirically significant: under the pressure of the evidence presented by Lavoisier and other 'modern' chemists, it became more and more difficult to maintain a set of hypotheses (H:) which credibly linked the various uses of 'phlogiston'[121]. Recognizing the gap between (mere) entrenchment and certifiable empirical significance allows us to consider as scientific activity the proposal and subsequent negotiation of explanations which we now know to be empirically insignificant[122].

V.

Our initial characterization of scientific activity stated that scientists were concerned with perceived regularities of nature for the purposes of explanation and prediction. Probing the diachronic and synchronic dimensions of 'explanation', we have now arrived at a more elaborate statement. Scientists propose explanations which are abductively successful and fulfill minimal conditions (C:)

and (D:); they then adopt a problem-raising attitude towards them, seek further reduction sentences for the predicates intoduced by the general laws of the explanations, and from then on negotiate whether those laws are independently testable and certifiably significant, i.e. whether the terms occuring in the laws are non-problematic. One might add that independently testable laws and non-problematic terms are not any longer subject of scientific concern but simply used in the course of further inquiry into other matters – unless, of course, new evidence calls their status into question.

Though this is still a pretty sketchy characterization, it has already begun to generate 'alternatives in science'. Once a proposed explanation is abductively successful and meets conditions (C:) and (D:), scientists have to choose between a problem- or mystery-raising attitude towards that explanation. Since only the choice of the former ensures the entertainability of an explanation and makes it a 'scientific explanation', insofar as it now becomes the subject of further scientific negotiations, the choice in itself does not present much of a problem here. But as was pointed out before (see above, page 98), it may become an issue whether demonstrably and regardless of professed intent a given explanation is actually treated in a problem- rather than mystery-raising manner. Thus, that choice will present a negotiable problem as it becomes the subject of controversy. The other set of alternatives arises when a second reduction sentence has been introduced for a term which is used in the context of an explanation: is that term still problematic or already non-problematic? If we decide that it is unproblematic, we thereby declare it to be certifiably significant, we transform the extant reduction sentences into at least a preliminary definition and a set of meaning postulates, we will view certain unanticipated observations as anomalies rather than falsifications, and our explanatory laws will be independently testable. However, making that momentous decision involves a possibly tenuous judgement on the truth of an empirical hypothesis ensuring the co-extensiveness of the term in question in all the contexts of its theoretical use. The issue of whether a term is problematic or non-problematic may therefore resurface at any time during the life-span of a theoretical term. And of course, there is another set of alternatives cutting across the others and lending further dimension to them: considering a term non-problematic is tantamount to accepting it, considering it as mystery-raising is tantamount to its rejection as a scientific term, yet in between we all along face the alternative of (tentative or final) rejection and acceptance.

I trust that the discussion so far motivated sufficiently well these sets of alternatives in science. However, there is a conspicuous gap in our updated characterization of 'scientific activity' where it is tacitly assumed that the adoption of the problem-raising attitude automatically entails seeking and finding further reduction sentences. Closer scrutiny of this tacit assumption shows it to be not quite true: making it involves yet another choice.

Indeed, if we look back at our discussion of how to supplement minimal conditions (C:) and (D:) of empirical significance, we discover an analogous gap. At one extreme, we can say that explanations which are not abductively successful, or which do not meet conditions (C:) and (D:), or towards which we adopt a mystery-raising attitude are not empirically significant and not entertainable as scientific explanations - and inversely, that a subset of 'scientific explanations' among the set of all 'explanations' is formed as soon as one adopts a problem-raising attitude towards an explanation which is abductively successful and meets conditions (C:) and (D:). At the other extreme, we can say that independently testable laws and non-problematic terms are certifiably empirically significant. And for some of the middle ground between these extremes, we can say that we are actively engaged in a negotiation on whether theoretical terms (entrenched by two or more reduction sentences) are problematic or non-problematic. But what about an entertainable 'scientific explanation' in the form of (I) which contains an explanatory term which is purely dispositional? We have adopted a certain attitude towards it which expresses, so to speak, the good will to find further reduction sentences for that predicate, and to *make* the term certifiably significant at some point in time. Yet, how long can can one thrive on good will alone, i.e. how long should one be willing to wait for the production of at least one other reduction sentence, and under which conditions should we revoke our tentative acceptance of (I)? And even more pointedly: does one have to ever find an explanation of the form (Ib) as a supplement to an explanation of form (I) in order to justify the acceptance of (I), be it as a true, problem-raising, entertainable, scientific explanation or as a very promising potential answer to a problem?

For instance, the early history of 'universal gravitation' revolves *inter alia* around the argument that one should adopt towards it a problem- and not a mystery-raising attitude (treating it as a *vera causa*), and that at the same time one should not expect to find an explanation (Ib) of why matter exerts gravitational force. The famous verdict '*hypotheses non fingo*' does not in principle rule out that an explanation of type (Ib) might be possible at some future time. However, it effecticely rules out the search for such an explanation for the present. Thus, 'universal gravitation' became a powerful explanatory tool without first becoming the subject of microstructural exploration[123]. The same situation holds whenever science reaches a fundamental level of explanation which, for the time being, is not considered a feasible subject of further investigation. Thus, we would have to distinguish among problem-raising attitudes as we assess the promise and explanatory value of an entertainable hypothesis:

- firstly, there is a problem-researching attitude which considers the disposition-predicate as a placeholder-term that is subject of a (research) programme that will ultimately render it non-problematic[124]; and

- secondly, there is a problem-deferring attitude, which considers the disposition-predicate a stopgap-device[125], further research of which is considered not feasible at or undesirable for the time being[126].

By choosing the placeholder-interpretation over the stopgap-interpretation for disposition predicates, one makes a commitment to a set of expectations concerning the future development of science in respect to the predicate in question. Accordingly, one also provides a statement, if only a weak one, of the conditions under which the choice should be reconsidered. If, within a certain period of time, the immediate resarch perspectives do not nearly live up to the promise of the hypothesis, and if the purely explanatory value of the hypothesis does not warrant its acceptance as problem-deferring, it is probably most opportune to altogether reject the hypothesis as mystery-raising.

By adding this third set of alternatives to the previous ones, we conclude this chapter by presenting the following 'boardgame-image' of science which serves to represent 'scientific activity' in its synchronic and diachronic aspects as culled from Hempel-Oppenheim, Peirce, and Carnap:

The Boardgame-Image of Science

Ab- duction	Discovery: explanatory hypothesis is proposed	
R O U T	routine determination of abductive success: the hypothesis is an 'explanation' and candidate for 'scientific explanation' iff it satisfies conditions (R1) and (R2)	
De- duction I N E	routine determination of 'prima facie entertainability': the hypothesis has prima facie empirical content and satisfies (R3) minimally iff conditions (C:) and (D:) are satisfied	
Alt I: O N G O I	problem-raising attitude the hypothesis is entertainable, a 'scientific explanation', candidate for 'true, scientific explanation'	mystery-raising attitude thus treated, the explanatory hypothesis is empirically insignificant, i.e. rejected as a 'scientific explanation'
Alt II: N G P R O C E S S O F C O L L E C T I V E	problem-researching attitude the placeholder-interpretation of disposition terms consists of the resolve to find further laws which are true for (at least) all those substances to which the (pure) disposition of the explanatory hypothesis is attributed, i.e. to find more reduction sentences for that predicate (the tentative acceptance as a 'true scientific explanation' is implied) ... time is passing ... Loop: adopt problem-deferring attitude after all? Loop: should hypothesis be rejected as mystery-raising after all? an appropriate law (further reduction sentence) is discovered: the (pure) disposition term becomes entrenched, i.e. a theoretical term	problem-deferring attitude the stopgap-interpretation of disposition terms rests on the judgment that, at the present time, search for further reduction sentences is either not feasible or undesirable Loop: adopt problem-researching attitude after all? Loop: is treatment after all mystery-raising? Alt: accept/reject as problem-deferring 'true scientific explanation'
Alt III: N E G O T I A T I O N O N (R 3)	the theoretical term is entrenched so that it is non-problematic condition (D‡) is satisfied iff hypotheses (H:) are held to be true, then: the general laws in which the predicate occurs are independently testable in respect to it and are certifiably significant, i.e. the fullest specification of (R3) is satisfied and the hypothesis is thus accepted as a 'true scientific explanation': it becomes part of background knowledge - unless it is challenged, leading into the Loop: a problematic entrenchment after all?	(still-)problematic the search for further laws continues; - if hypotheses (H:) are held to be plausible, Loop: is the predicate non-problematic after all? Alt: accept/reject as problematic theoretical predicate - if hypotheses (H:) appear untenable, Loop: adopt problem-deferring attitude after all? Loop: should the hypotheses in which the term occurs be rejected as mystery-raising?

Like all flowcharts, the 'boardgame-image' of science provides a representational device for certain sequences rather than an implicit hypothesis concerning strict historical chronology. Due to a number of factors, many layers of the sequences represented in the flowchart will partially overlap in any given period of time.

- Often, various scientists can probably be found at different positions in the game. Thus, there is simultaneous argument (and search for evidence) at various levels of the chart.
- Scientific theories often contain several key-concepts, any number of which may still be problematic (or become problematic again). In respect to certain terms of a single theory, we might accordingly find ourselves at different stages of the game.
- The chart itself contains a number of loops, i.e. it designates places at which upon certain judgements discussion of a predicate or the property designated by it may revert to any earlier stage.

Each of these three factors may lead to the introduction of new relevant evidence which will affect other scientists, other theoretical terms, and previously made decisions.

As a device for representing alternatives and sequences of decision-making at the core of scientific activity, the flowchart presupposes only a minimum of normative considerations: the routine decisions concerning (R1) and (R2) (abductive success) and (C:) and (D:) express normative constraints on 'scientific explanation' as well as the demand that the problem-raising attitude should be preferred over the mystery-raising one. Beyond that, the flowchart functions neither pre- not proscriptively. For instance, it is not implied that a predicate should be rendered unproblematic. Whether it should be or not, will in any given case depend on matters such as its promise and explanatory value, the availability of evidence, feasibility-judgements, demands for information, the competitive pressure exerted by rivaling theories, degrees of confirmation or corroboration, simplicity, the frequency and character of anomalies or falsificatory instances etc. Judgement upon these matters may render a hypothesis that is interpreted in a problem-deferring manner as a perfectly adequate true scientific explanation[127].

While systematic issues are at stake at any point in the flowchart, the temporal dimension becomes prominent as the scientists enter into the phase of deduction and conclude their routine-assessments of abductive success and minimal empirical significance[128]. The sheer passage of time weighs heavily during the phase between the choice of a problem-raising attitude and the presentation of a general law which serves to theoretically entrench an explanatory concept. One might expect that the collective deliberation of alternatives in science *qua* alternative attitudes to be adopted towards a disposition term is especially lively and explicit during this period of active waiting. The history of a theory which for a long time lingered in this (lively) position of relative stagnation may therefore provide a splendid illustration of the heuristic and taxonomic powers of our theory of alternatives in science.

Notes to Chapter 2

1 According to the standard for the progressivity of methodologies which we adopted from Lakatos in the previous chapter, 'maximally inclusive' and 'minimally pre-emptive' are relative to extant theories of science: both constraints lead to a maximally internal reconstruction of the history of science.

2 Due to the constructive and as such somewhat arbitrary character of that task, it is only appropriate that - by using words like 'we' and 'us' - I explicitly invite the reader to join me in what can only work as a shared effort.

3 These formulations are controversial among contemporary philosophers of science. Not all criticism of Hempel-Oppenheim's explication or Carnap's programme can be discussed here. However, an attempt will be made to show that in many cases the grounds for dissatisfaction serve to underline rather than subvert the well-suitedness of these proposals for the purpose at hand. From the point of view of (PI) and a research programme that is aiming only for a conceptual framework for the methodological discussions of scientists, competing philosophies of science frequently appear complementary rather than contradictory.

4 This kind of genetic approach towards a formulation of alternatives in science was also taken by Böhme et.al. [1972], p. 312. They maintain that alternatives in science present themselves at those junctures in the rational reconstruction of the history of science where methodological theories of scientific progress fail to uniquely determine the future course of theory development. That is, alternatives in science arise as the (purely) methodological pursuit of scientific activity generates such situations. On this account, the study of alternatives can then provide the "missing link" between the "internal relations [Binnenstruktur] and the social relations [Sozialstruktur] of science".

5 Compare e.g. Popper [1975], p. 191:

> it seems that when we speak of science we do feel, more or less clearly, that there is something characteristic of scientific activity; and since scientific activity looks pretty much like a rational activity, and since a rational activity must have some aim, the attempt to describe the aim of science may not

> be entirely futile. I suggest that it is the aim of science to find *satisfactory explanations*, of whatever strikes us as being in need of explanation.

6 See Hempel and Oppenheim [1965], p. 249:

> an explanation of a particular event is not fully adequate unless its explanans, if taken account of in time, could have served as a basis for predicting the event in question.

7 Hempel and Oppenheim [1965], pp. 247f.

8 Hempel [1965], p. 412.

9 Since the deductive-nomological model of explanation has come under attack from various sides (see e.g. Levi [1969], Van Fraassen [1980], Cartwright [1983]), I should make very clear what that representational device represents: it designates artifacts which are the object of scientific negotiation. That is, regardless of how one has arrived at an explanation and what further fitting of the facts to the theory is required, the explanation will be presented in a form that is representable within the H-O model. Indeed, it would be most misleading to draw any inference from the form in which explanations are presented to the way in which explanations are produced. Cartwright [1983], p. 104 rightly objects to this sort of interpretation of the deductive-nomological model:

> It is never strict deduction that takes you from the fundamental equations at the beginning to the phenomenological laws at the end.

Cartwright's own construal of 'explanation' also includes a deductive relation of explanans to explanandum. It differs from the deductive-nomological account insofar as it does not admit a facile interpretation that places the general law in the explanans at the beginning of some temporal development, the explained facts at the end. See e.g. p. 152:

> To explain a phenomenon is to find a model that fits into the basic framework of the theory and that thus allows us to derive analogues for the messy and complicated phenomenological laws which are true of it.

In contrast to the position which Cartwright so convincingly attacks, I do not propose that there is a diachronic analogue to the deductive structure of an explanation (though there is a diachronic analogue to the criteria of adequacy which those explanations are to meet). Further, I do not use the model to embellish a realist programme in the philosophy of science, i.e. I allow for all kinds of modelling and fitting of facts to theories that would lead to constant modifications of the explanation under consideration. - The term 'explanation' as used in this investigation, can be replaced at any time by 'simulacrum account' (Cartwright), 'semantically adequate account' (Van Fraassen), or 'application of a rule for material inference' (Levi). (Also: I believe that all the problems pertaining to deductive-nomological explanation also pertain to inductive-statistical explanation but not *vice versa*.)

10 Hempel [1965], pp. 412f.

11 Hempel [1965], p. 249.

12 One might also say that (R3) selects a subset of 'empirical', 'informative', 'non-vacuous', or 'good' explanations. I have chosen the label 'scientific explanation' for two reasons. Firstly, it captures the intentions of the H-O model of *scientific* explanation. Secondly, I will later argue that scientists who debate whether a given proposal is properly 'scientific' or not, are actually attempting to apply (R3) (or a version thereof) to that proposal.

13 Or, for a more drastic example, consider a child's question, "why is *x* so?", and the parent's answer, "because this is how things are", or, "because this is how it has always been": the answer, in both cases, can be reconstructed in terms of (R1) and (R2) - serving as *heuristic* devices the two conditions are satisfied by *any* conjectural statement of a general law to which the parent has tacitly and elliptically appealed.

14 To safeguard against misunderstanding, I should emphasize that the expediency of doing so is not at issue here. After all, the model is not intended to represent the *beliefs* of persons who advance an explanation. For certain contexts, a more accurately descriptive model is desirable. Thus, we find Arthur Danto on historical explanation and Donald Davidson on the explanation of actions adopting different models - only after acknowledging the applicability in principle of the H-O model. See Danto [1965], pp. 201-232; Davidson [1980], pp. 261-275.

15 From here on in, 'explanation' shall be used only in reference to contexts in which an explanandum is to hold *because* an explanans holds.

16 It was pointed out (but not developed) e.g. by Scheffler [1981], p. 128.

17 In the following, I heavily draw upon Peirce's theory of abduction (after 1901) as restated by Fann [1970]. -- The distinction between abduction and induction may involve an unintuitive notion of 'induction', delegating to 'abduction' what is frequently considered paradigmatically inductive, namely the process of generalization. That Peirce himself held different views on this distinction, heightens the confusion. Contrasting the 'early' and the 'present' (1901) views, Fann writes on pp. 33f.:

> According to the early view both abduction and induction are 'synthetic' in the sense that something not implied in the premises is contained in the conclusion. [...] Induction is reasoning from particulars to a general law; abduction, reasoning from effects to cause. The former classifies while the latter explains. On Peirce's present view *any* synthetic proposition, whether it is a nonobservable entity or a generalization (so-called), in so far as it is for the first time entertained as possibly true, it is an hypothesis arrived at by abduction. [...] On the early view this last example of an abduction would have been called a 'generalization' which would only be the result of induction. On the present view such generalization is *suggested* by abduction and only *confirmed* by induction.

Here, only the later ('present') view is appealed to.

18 Peirce [1958-60], 5.171 (comp. also 5.590).

19 Ibid.

20 Peirce [1958-60], 5.188f.

21 See also Peirce [1958-60], 7.202 and especially 5.146:

> Abductive and Inductive reasoning are utterly irreducible, either to the other or to Deduction, or Deduction to either of them, yet the only *rationale* of these methods is essentially Deductive or Necessary.

Clearly, Peirce's theory of abduction is far more complex than what is considered here. And yet, for example perceptual judgements - which according to Peirce are limiting cases of abduction - also conform to (R1) and (R2), see e.g. 5.150 and Apel [1975], pp. 208 and 300ff.

22 'Abductive awe' designates a certain kind of response by scientists to a suggested explanation. The presence of abductive awe thus reveals the perceived

initial plausibility of that proposal. The connection between perceived initial plausibility and economical features such as caution or boldness, breadth, and simplicity is more than tenuous: the latter considerations enter into an evaluation of the anticipated future performance and usefulness of the theory, i.e. its promise. Initial plausibility as it is revealed in abductive awe, on the other hand, is a function of the degree of difficulty attributed to the task of coming up with *any* unified account for a given range of phenomena. Abductive awe looks backwards at an intellectual accomplishment while evaluation of promise is prospective.

23 Lichtenberg [1975], p. 261 (J 1416):

> Ist es ein Traum, so ist es der grösste und erhabenste der je ist geträumt worden, und womit wir eine Lücke in unseren Büchern ausfüllen können, die nur durch einen Traum ausgefüllt werden kann, und ist der Traum in sich zusammenhängend, entfernt man sich nie von den Vorschriften einer richtigen Analogie, so kann er ja die Wahrheit selbst sein oder ihre Stelle vertretten.

24 Lakatos [1978c], pp. 170-191.

25 This verdict is tentative since it does not imply that stronger answers are impossible. On the strategy adopted here, it may well turn out that scientists uniformly adopt the same way of supplementing conditions for 'scientific' explanation. Upon the results of future research, we might inch towards stronger claims on the connection of methodology and scientific practice.

26 Though most will agree that no other approach has been proposed, it may not be obvious why (i) and (ii) actually present an exclusive and exhaustive disjunction of possible approaches. I suggest that - at this point - 'test' be considered a broad enough notion to include any mediated or indirect method for determining 'empirical content', while (i) proposes the only direct and unmediated methods. -- The difference between (i) and (ii) was a live issue e.g. in Carnap [1936/37]. On p. 423 Carnap argued against Schlick that a sentence

> S_1 is confirmable not because of the logical possibility of the fact described in S_1, but because of the physical possibility of the process of confirmation.

27 Note that this preliminary formulation entails a falsifiability requirement: by modus tollens, there is

retransmission of falsity from the test-statement to the conjunction of premises from which it has been deduced. However, as of yet we are far from making the additional methodological demand that 'falsification' should lead to the abandonment of any of those premises.

28 As a limiting case the explanatory hypothesis may itself be a test-statement. In this case, the testable statement is deduced from the explanatory proposition since it is true that (P => P).

29 In a very similar fashion, Peirce (in 5.197f., see also 7.203) motivates the transition from the process of 'abduction' to 'deduction' in terms of how 'good' explanatory hypotheses can be designated. Indeed, this inquiry into scientific decision-making seeks to revitalize Peirce's interest in deduction, the 'economy of research' and his conviction that "What is good abduction?" is "the question of Pragmatism" (5.197).

30 For the present purposes we can safely bypass Peirce's theory of signs by substituting 'model' for 'icon'.

31 Peirce [1958-60], 5.161f.

32 Peirce [1958-60], 7.206.

33 To be sure, the notion of 'acceptance' always involves the notion of truth. More precisely, then, the different acceptabilities should be stated as 'to find true that the sentences are true', 'to find true that the sentences yield the strongest potential explanation', etc. The notion of acceptance as employed here (and as employed by Peirce) thus coincides with 'formation of belief'. -- Condition (R4) as stated by Hempel and Oppenheim does not reflect a differentiation of acceptabilities and entertainabilities. Thus, our theory of alternatives in science employs a more general equivalent to (R4) rather than (R4) itself.

34 For this reason, the following will make extensive use of that text.

35 Carnap [1956], p. 53. To be sure, it seems that in that later paper Carnap significantly modifies the view he held in *Testability and Meaning*. However, it will become apparent that the later article expresses a shift of perspective rather than of opinion.

36 Carnap [1936/37], p. 452.

[37] Carnap [1936/37], pp. 468-471 suggests that just one observational term, e.g. 'bright' or 'solid' may constitute a sufficient observational basis for the language of contemporary physics.

[38] Carnap [1936/37], pp. 454f. and compare p. 448 (keeping in mind that the sentence 'P(b)' used in the explication of 'observable' is an atomic sentence.):

> It should be noticed that the term 'atomic sentence', as here defined, is not at a.1 understood to refer to ultimate facts. Our theory does not assume anything like ultimate facts. It is a matter of convention which predicates are taken as primitive predicates of a certain language L; and hence likewise, which predicates are taken as atomic predicates and which sentences as atomic sentences.

Since 'conventional' can be construed as 'relative to a reseach programme', admission of conventionality is not at all tantamount to admission of arbitrariness into scientific discourse.

[39] Carnap [1936/37], p. 455.

[40] Carnap [1936/37], p. 443.

[41] Carnap [1936/37], p. 439.

[42] I.e. if the conditional is true it expresses a structural property of nature, if it is proposed as possibly true then it designates a presumed property of nature.

[43] Since not all generalized conditionals are equal, there is not always a need to name the structural properties expressed by them. (i) 'If x is immersed in water, then x dissolves' is not a physical truth. It is true only for certain objects. Conditionals like this are often used to introduce names for properties (e.g. 'soluble'), since the availability of these names serves the useful function of classifying physical objects into those which have that property and those which do not. (ii) 'If x is in the gravitational field of the earth, then x falls freely towards the center of the earth' is also not a physical truth, it is not true e.g. for gases. Like (i), it introduces a property that serves to distinguish all objects into those that have it and those that do not. However, if we want to determine whether something has that property or not, we need not ordinarily question whether the antecedent condition is fulfilled, since, for all practical purposes within terrestrial physics, it is always fulfilled. Here, the

consequent suffices for defining that property: this is the limiting case mentioned above in which (conditional) physical reduction becomes (unconditional) definition. (iii) 'If x is a swan, then x is white' is (presumably) a physical truth. It designates a property of nature. There is little need in cases like these to introduce a name for the dispositional property shared by everything that it is white under the condition that it is a swan. (iv) 'If x consists of matter, then it exerts gravitational force' combines (ii) and (iii): the antecedent is always satisfied and the conditional itself a physical truth. Sometimes, conditionals of this sort are called 'laws of nature'. Again, the properties of nature tacitly introduced by such conditionals do not need to be named.

44 It is presupposed that this definition meets the requirement that all terms occuring in the S- and R-statements belong to the old language, i.e. that they either belong to the observational language or have previously been introduced by reduction (or a chain of reductions) to the observational basis.

45 If this line of criticism is not sufficiently compelling in the case of 'soluble', how about defining e.g. 'electrical charge' as 'deviation of a magnetic needle'?

46 This view of scientific language has been elaborated and carefully defended by Shapere, see especially his [1984], p. xxxv:

> it is the concept of "reasoning" that is fundamental to the interpretation of science, and the interpretation of scientific language is derivative from that.

47 Carnap [1936/37], p. 440. - This becomes even more obvious when we transform the conditional so that the definiens of D reads -(S & -R).

48 This generalization does not encompass those cases where the antecedent of the conditional (S) happens to be always satisfied, whether trivially (if it is analytic) or contingently so (compare note 43 above). -- Gethmann [1980], p. 24 suggests that the paradoxical result of 'soluble wood' stems from the mistaken choice of standard logic in the analysis of the definition. He argues that it would not arise in the framework of a non-standard logic ("Minimallogik") which permits us to draw the conclusion --R from the premises -S and (S => R), but which does not permit us to go from --R to R. Using this non-standard logic, the attribution of D is therefore not warranted on the basis of -S. However, the issue before

us is not how certain conceptual problems concerning the introduction of disposition predicates can be bypassed or solved. The issue is rather how to best represent the problems upon which scientists deliberate. I would suggest that scientific problems are most adequately framed within standard logic, i.e. that we should impute to scientists that they adhere to standard logic. This imputation seems quite justified in view of the ongoing scientific negotiation concerning the introduction of disposition predicates that are a bit more controversial than 'soluble'. But again, the imputation is self-correcting in that it can lead to the discovery of greater sophistication in scientific ways of framing problems. (This argument applies *mutatis mutandis* to any other attempt at solving the problems associated with disposition predicates by altering the language in which they are commonly introduced, as e.g. the substitution of a 'subjunctive conditional' for the conditional by Fetzer [1981], pp. 36ff.)

49 Carnap [1936/37], pp. 442f. defines 'bilateral reduction sentence' as a special case of a 'reduction pair', where:

$$\text{reduction sentence: } S \Rightarrow (R \Rightarrow D)$$
$$\text{reduction pair: } S_1 \Rightarrow (R_1 \Rightarrow D)$$
$$S_2 \Rightarrow (R_2 \Rightarrow -D)$$

50 Carnap [1936/37], pp. 440f. - Again, it is presumed that S and R can be expressed in the old language. - From here on, following Carnap's convention (see p. 434), I will drop the indices unless they become particularly relevant.

51 Carnap [1936/37], p. 458.

52 This sentence represents the factual content of the bilateral reduction sentence and is therefore called 'representative sentence', see Carnap [1936/37], p. 451.

53 A more precise formulation should eliminate the redundancies in the following.

54 Again, a methodological connection between 'falsification' and 'rejection' is not assumed here.

55 Carnap was cited above as saying that the possibility of introduction by law is "very important for science, but so far not sufficiently noticed" (Carnap [1936/37], p. 443). Yet, he himself did not elaborate the significance of this. The clear differentiation of 'test for' and 'test of' on the one hand, 'use' (in general

laws, explanation and prediction) and 'mention' (in conditional definitions, bilateral reduction sentences) on the other hand, provides a first clue: after all, more often than not, reduction sentences remain tacit in scientific discourse while the terms are introduced as they are used, i.e. in general laws.

56 I believe Ryle [1949] was the first to emphasize the use of disposition terms in explanatory contexts. See e.g. pp. 124f.:

> Dispositional statements about particular things and persons are also like law statements in the fact that we use them in a partly similar way. They apply to, or they are satisfied by, the actions, reactions and states of the object; they are inference-tickets, which license us to predict, retrodict, explain and modify these actions, reactions and states. [...] Dispositional statements are neither reports of observed or observable states of affairs, nor yet reports of unobserved or unobservable states of affairs. They narrate no incidents.

This account has raised the ire of so-called scientific realists like Mellor [1978]. While Ryle's construal of dispositions accords well with their introduction by reduction-sentences, Mellor's attempt to construe dispositions as real properties leads to some constraints on the introduction of dispositions which are not met by Carnap's method. On closer scrutiny, however, Mellor's account differs from Carnap's only in one respect: he does not want to admit the inference [(S & R) => D] (Mellor [1978], pp. 64f. and 71). Mellor does not provide an argument for this counter-intuitive exclusion. This might exemplify that in the debate on dispositions (just as anywhere else) the debate on realism versus e.g. conventionalism concerns a pseudo-problem in the philosophy of science.

57 Not even with inductive rules that will warrant our acceptance as true of the proposition that some proposition is entertainable or a 'scientific explanation'. For, as we shall see, the minimal constraints elaborated here do not suffice when it comes to deciding whether or not to entertain a hypothesis.

58 That 'testability' and not 'testedness' should be required, has always been taken as self-evident, even by Carnap. Compare also Hempel and Oppenheim's formulation of (R3) in their [1965], p. 248:

> The explanans must have empirical content; i.e., it must be capable, at least in

principle, of test by experimentation or observation.

While I do not wish to question the plausibility of this very basic assumption, it should be noted that it leads to a self-exemplifying feature of the philosophy of science: Schnädelbach [1971] pointed out that 'observable' and 'testable' (like 'confirmable', 'falsifiable', etc.) are themselves disposition-terms. I believe that here it is crucial to follow Lakatos' intuitions on the relation of philosophy and history of science: Historical explanations brought forth by the science of science should meet our minimal requirements on 'scientific explanation' and on the use of dispositional predicates in the context of such explanations. Accordingly, the key dispositional concepts introduced in this chapter can easily be brought into the form of bilateral reduction sentences satisfying (C:) and (D:): just consider Carnap's (conditional) definition of 'observable' as quoted on p. 68; (C:), transformed into a reduction sentence, introduces 'testable'; and the answer to (D) will introduce 'entertainability' (given that a hypothesis is entertainable if and only if it fulfills minimum conditions of empirical significance). Of course, this does not render these predicates unproblematic. On the contrary, it establishes an uncanny continuity between science and science of science. [Incidentally, one philosopher very aware of that continuity is Nelson Goodman, who introduces the disposition term 'projectible' in his [1973], pp. 86f.]

59 For Carnap's argument to that effect see his [1936/37], pp. 25f. and 28f.

60 Carnap's argument for 'observability' over 'realizability' is stated in his [1936/37] on p. 462 (in his terminology, this is the decision for 'observability' over 'testability'). He presents a different sort of example: science may usefully distinguish people who are disposed to get a certain disease from people who are not so disposed simply by looking at those who at some time contract that disease and for whom we find that certain other properties hold - regardless of whether it is possible to experimentally test for that disposition (whether we can induce or produce the disease).

61 This formulation has to be suitably amended in order to cover the occurence of several predicates D in a proposition: that proposition is empirically significant iff it is testable in respect to each occuring D. -- For Carnap's statement of the decisions leading up to (D), see his [1936/37], pp. 33ff. - As an example of how liberal these conditions are, Carnap presents the

following sentence as an empirically significant proposition:

> If all minds (or: living beings) should disappear from the universe, the stars would still go on in their courses. (p. 37)

62 Of course, in order to interpret a lack of consensus as indicative of a fundamental problem, I have to assume not only that there is some agreement on this lack of consensus, but also that it should be ascribed to substantive problems in each of the competing approaches and not to stubborness on the part of certain factions.

63 For an overview of the criticisms (in respect to disposition predicates) see Essler and Trapp [1978].

64 See above, p. 54.

65 Pap [1963], p. 570, Mellor [1978].

66 Essler and Trapp [1978] suggest a variation on Carnap's method of introduction. Their reduction sentence explicitly expresses the belief that response R will unfailingly occur:
$(x)\ [(Et)\langle S_{xt} \Rightarrow (D_x \Leftrightarrow (t)[S_{xt} \rightarrow R_{xt}])\rangle]$

67 Thus, the bilateral reduction sentence for 'having positive electric charge' may attribute that property immutably to all pieces of iron in a state that is temporally delimited by t_1 and t_2, where at t_1 one sort of operation is performed on the object (so that it is charged) and at t_2 another operation (so that it loses its charge). This use of 'immutable' may be counter-intuitive but it is unavoidable: if we did not construe 'immutability' for objects and states within longer or shorter time-intervals, every property would be transient and the intuitive force of the distinction would be altogether lost.

68 Though this cannot be substantiated here, I believe this to be true also for explanations in which the general law is statistical or probabilistic.

69 The fact that many composers do not strictly follow the rules of harmony does not make these rules a less valid tool for researching the methods of those composers. Here too, the ways in which they deviate from these rules provide the master-clue for an assessment of their methods, significance, or style.

70 Isaac Levi [1980], p. 244 emphatically underlines these limitations of purely philosophical reasoning:

> We improve our understanding of predicates such as 'is compelled to attract iron filings placed nearby' by studying magnetic theory, and not by studying possible worlds or any other armchair semantics.

This can be nicely juxtaposed with Goodman [1973], pp. 46ff. which equally emphatically maintains that disposition predicates should be defined:

> Philosophy, to my way of thinking, has rather the function of explicating scientific - and everyday - language than of depicting scientific or everyday procedure. [...] The argument that we do better to refrain from defining a term in explanatory discourse unless that term is customarily defined by scientists or laymen is like the argument that philosophy ought not to be coherent unless the reality it describes is coherent. One might as well argue that philosophy should not be written in English because the world is not written in English. There is no positive virtue in not defining disposition terms. (pp. 47f)

Without wishing to discount the intuitive force of Goodman's argument, I should point out that we arrive at a position contrary to his not because we see great virtue in not defining disposition terms, but because of the apparent impossibility of defining them in agreement with scientific practice (and it is not at all clear how Goodman has succeeded in giving a definition).

71 This view has been expressed in a variety of ways. The most forceful statements can be found when the debate turns on the question whether dispositional explanation is *causal* explanation (where it usually remains tacit that only causal explanation is properly scientific). Focussing points of this debate are statements like this passage from Armstrong [1969], p. 26:

> it is linguistically proper to *identify* brittleness with that state of the brittle thing that, if the object is struck, causes it to break. It is linguistically proper, I assert, to say that brittleness *is* a certain sort of bonding of the molecules of the brittle object. In this I think I am following the way scientists are prepared to speak.

This statement provides a nice example of 'armchair semantics' (see the preceding note). For further contributions to the exchange surrounding Armstrong's assertion, see Squires [1968], Coder [1969], Stevenson [1969], but also Ringen [1982]. Of these contributions, only Squires [1968] addresses the issue along the lines of the following attempt at clarifying the issue by introducing a presumably unacceptable dispositional explanation.

72 This explanation exemplifies rather nicely Ryle's notion that disposition predicates function as inference-tickets (see above, note 56). It should be noted that the disposition predicate could be eliminated from this explanation: the explanatory function of the disposition predicate is not that of a law. Rather, its occurence in a law activates a rule which licenses the inference to the explanandum. Compare Levi [1969], pp. 300f. which recommends this approach to dispositional explanation (and a modified approach to statistical explanation).

73 Here, a variation on the reduction sentence might do a more compelling job:
$(x)(y)\ (S_{xy} \Rightarrow [D_x \Leftrightarrow R_y])$, where S: 'x is introduced into y' and R: 'y falls asleep'.

74 No detailed argument for this should be needed here. Indeed, given the symmetry of explanation and prediction, the testability of the general laws in (I) and (II) obtains trivially: R_a is not only the explanandum but also constitutes the truth-condition in a test of the attribution of D by the general law, i.e. it tests the general law.

75 Moliere [1950], p. 78.

76 The use of 'seemingly uninformative' and 'apparently futile' indicates that neither (I) nor (II) are completely uninformative: both are generalizations asserting regularities of nature, and both specify a causal agent or, rather, a substance in which we can presumably locate the causal processes at work. Compare Squires [1968], p. 47:

> Thus it would be wiser to say, not that a dispositional explanation gives the cause of an event, as that it shows where the cause is to be found. It is rather like saying that the weather is responsible for the good crops. This rules out certain explanations, such as the richness of the soil or the special breed of corn. But it only indicates the area in which to look for a cause. [...] Dispositional explanations leave room for certain types of causal explanation rather than provide them.

For a more comprehensive statement of this position, see Levi and Morgenbesser [1978], pp. 401f.:

> Disposition predicates, like *ceteris paribus* clauses, function as placeholders for

predicates specifying conditions in generalized statements. But they are not simply place-holders and differ from *ceteris paribus* clauses in a number of respects: *Ceteris paribus* clauses entail no commitment as to the kinds of predicates to be employed in replacing them; disposition predicates do. [...] Thus, if we are told that objects break when tapped lightly *ceteris paribus*, we are told very little about how the blatant deficiencies of this generalization as a law can be eliminated, save that Hemepl-Oppenheim requirements are to be met. We do not know what kinds of predicates are to be used [...] For all we know they may refer to observable conditions in the environment, observable features of the windows or the micro-structures of the windows. On the other hand, if we are told that fragile objects break when tapped lightly, we assume that if we are to improve or replace the generalization we should investigate the micro-structure of fragile objects.

77 This distinction (and its ensueing differentiation) is taken from Levi [1967], p. 194.

78 It should be noted that presumed goal-directedness and teleological thinking is a problem here only insofar as it is entailed by the 'mystery-raising' attitude: there is no intrinsic connection between teleology and the explanation as formulated in (II). And indeed, while contemporary pharmaceutical scientists may not use 'dormitive power', they do use equivalent predicates to divide all substances into 'tranquilizers' and 'non-tranquilizers' etc. Accordingly, our common-parlance explanation "person *a* fell asleep because she took a sleeping-pill" appeals precisely to the kind of general law as the one stated in (II). [Incidentally, we find that by using mystery-raising explanations in order to underline the difficulties posed by disposition terms, philosophers appeal to an attitude which has become obsolete with the demise of Aristotelian science.]

79 For the following, I have consulted a few textbooks and dictionary articles on solutions and solubilities, especially Prausnitz [1982]. However, I have made extremely eclectic use of these articles - in order to prepare the ground for the following discussion, without attempting to give a reconstruction of chemical history.

80 Since historical accuracy is not the desideratum here, I am disregarding for simplicity's sake

'solubility' in respect to various solvents other than water, and also, degrees of solubility. That is: 'soluble' designates for now only the property of a substance to dissolve (vs. not-dissolve) in water.

81 Of course, we will have to ask our scientists later on (rather, they will ask themselves) whether these differences between sugar and salt can all be attributed to the property of 'electrolyte' versus 'non-electrolyte' solubility.

82 Prausnitz [1982], p. 1048. - Throughout, this article makes quite clear that we are dealing here largely with dispositional properties of substances. Consider, for example, the use of the verbs 'tend to' (p. 1047) and 'can' (p. 1048), and phrases like 'measure of the ability to' (p. 1048) and 'tendency to' (pp. 1048 and 1050).

83 Having two reduction sentences seems redundant here, since we have the limiting case that E and NE form an exclusive and exhaustive set of soluble substances, such that *not 'electrolytically soluble'* coincides with *'nonelectrolytically soluble'*.

84 Indeed, in most contexts we might find this an excessivley elaborate explanation and we might actually prefer simple-minded (I) as a straight answer to our question.

85 Of course, there are other ways of formalizing the results of the reported episode in the history of chemistry. For instance, we could introduce another disposition Di: 'x is disposed to dissociate into ions', where S_4: 'x is dissolved in water and is tested for ion-dissociation', and where R_4: 'x dissociates into ions'. Now, we could introduce another reduction sentence for NE and thus arrive at the following explanation (S: 'x is immersed in water'):

$$
\begin{array}{l}
(x)\ (Sg_x \Rightarrow -Di_x) \\
(x)\ (S_x \Rightarrow [NE_x \Leftrightarrow -Di_x]) \\
(x)\ (D_x \Leftrightarrow [E_x \vee NE_x]) \\
(x)\ (S_x \Rightarrow [D_x \Leftrightarrow R_x]) \\
Sg_a \\
S_a \\
\hline
R_a
\end{array}
$$

The following discussion pertains to this formalization (and many other possible scenarios) as well as to the one presented in the main body of the text.

86 For instance, on what may be the deepest level of energy considerations, the process of dissolution is tantamount to a decrease of the value resulting from enthalpy minus the product of absolute temperature and entropy. (Compare Prausnitz [1982], p. 1052.)

87 The development described here could be dubbed 'progressive disposition-shift'. After all, the sketchily constructed story of the successively more thoroughgoing explanation of R_a presents an analogue to Lakatos' theory of progressive problem-shifts in the growth of scientific knowledge: while he was focussing on the adjudication between rivaling theories, I am here primarily concerned with linear (non-revolutionary) growth of knowledge in normal science. In both cases, new problems are substituted for old ones. Indeed, Lakatos's methodology was most nearly adumbrated by Levi and Morgenbesser [1978] on disposition predicates (compare the extensive quote from their article in note 67 above). -- The progessive substitution of higher level disposition predicates for lower level ones (which indicates the ineliminability of diposition terms) has also been pointed out (though different lessons have been drawn from it) by Squires [1968], p. 45; Mellor [1978], p. 70; Popper [1968], p. 424; Armstrong [1978], pp. 417f. and 420; and Fisk [1978], pp. 207f.

88 Unfortunately, it is impossible to historically pursue this *topos* in the history of epistemology. Compare e.g. Duhem [1954], p. 218:

> We are thus led to the conclusion so clearly expressed by Claude Bernard: the sound experimental criticism of a hypothesis is subordinated to certain moral conditions; in order to estimate correctly the agreement of a physical theory with the facts, it is not enough to be a good mathematician and a skillful experimenter: one must also be an impartial and faithful judge.

Duhem himself ascribes his emphasis on common sense to the influence of Blaise Pascal. However, it may well be worthwhile to place into this context Kant's *Urteilskraft* as a mediator in the 'play of the cognitive faculties' (see section VI in the preface to the *Critique of Judgement*):

> our judgement makes it imperative upon us to proceed on the principle of the conformity of nature to our faculty of cognition, so far as that principle extends, without deciding - for the rule is not given to us by a determinant judgement - whether bounds are anywhere set to it or not. For while in respect of the rational

> employment of our cognitive faculty bounds may be definitely determined, in the empirical field no such determination of bounds is possible. (Kant [1911], pp. 28f.)

89 In the following, the term 'entrenchment' is used as a theoretical primitive, if only preliminarily. Nelson Goodman uses it as a criterion for 'projectibility' in his [1973], p. 101. To that end, he considers only the comparative entrenchment of terms occuring in rivaling theories. Without the luxury of such a restriction, the question arises just what it means for a term to be entrenched. After having dealt with this question, 'entrenchment' will designate the beginning of a particular stage in the development of a theory. That this stage is full of uncertainties (i.e. that to be entrenched can mean quite different things) jeopardizes Goodman's definition of 'projectibility' - and this is significant insofar as Goodman's transcendental argument for the definability of disposition terms depends on the successful definition of 'projectibility'.

90 Indeed, one may now recall that Peirce's notion of 'deduction' (see above, p. 64) entailed a tacit demand for the pragmatic asymmetry of explanation and prediction. At the time, it appeared as an apparent asymmetry of abduction and deduction.

91 Hempel and Oppenheim in Hempel [1965], p. 249. Compare also Hempel [1965], pp. 366f.

92 The underlying notion of progressive problem-shifts in the history of science can be crudely sketched as the following sort of progression: We regularly (not necessarily always) observe (S_a & R_a) and by manipulating and observing some more we find that unfailingly (S_a => R_a) obtains. Now we formulate the general law (x) (S_x => R_x) which explains R_a if initial condition S_a is fulfilled. We continue our investigation by looking for a Q_a such that we regularly observe (Q_a => [S_a => R_a]) which leads to the formulation of (x) (Q_x => [S_x => R_x]), which explains (S_a => R_a) if initial condition Q_a is fulfilled. Etc.etc. - This applies to all three of our explanations so far: (I), (II), and (Ib).

93 One can identify six evidential states vis-a-vis a general law of the form (x) (S_x => R_x): (i) (-S & -R); (ii) (S & R); (iii) (-S & R); (iv) Q; (v) -Q; (vi) (S & -R) [where Q is a consequence of our general law in conjunction with suitable background knowledge].

ways than one. This is no new idea. (Mellor [1978], p. 69)

On this account, a disposition *becomes* a real property as we learn more about it. Here, again, the presumed conflict between realism and conventionalism dissolves as one considers merely the dynamics in the growth of knpwledge. -- Finally, for a stimulating discussion of the problem of 'old evidence' (a critique of Glymour from a Bayesian perspective), see Garber [1983].

95 As in (Ib) which included the definition [D <=> (E v NE)].

96 Carnap [1936/37], pp. 444f.

97 I speak here of 'virtual definition' rather than 'definition' in order to acknowledge that a term virtually defined by a set of bilateral reduction sentences most probably was originally introduced not by definition but with the help of plain physical reduction sentences. And further, that there may still be independent means of attributing that term (or at least partially determining its meaning). While 'definition' often refers to the adoption of a convention that fixes the meaning of a predicate, 'virtual definition' merely characterizes a form of particularly tight entrenchment in a language. Thus, as was pointed out by Kuhn [1983], Newton's Second Law virtually defines 'force' and is nevertheless empirically contingent (it is an "axiom" *and* a "law"). The notions of entrenchment and progressive entrenchment (leading up to virtual definitions) thus shed light on Kuhn's claim that one cannot learn certain scientific terms without learning some others.

98 Carnap [1936/37], pp. 449f.

99 Newton [1958], p. 53 [p. 3081 in the original].

100 These three general laws are not quite on a par. The first two predict the outcome of a physical operation (measuring), the last the outcome of a mathematical operation on two measured values.

101 Again, we will find that properties characterized by equations are not always given a name. Presumably, this depends largely on whether the disposition is considered non-problematic or problem-raising. (Curious testimony to this lack of names is the occasional use of the mathematical structure of the equation for a name, for instance: 'inverse-squareness' taken as a property of physical states.)

102Note that the above-stated general laws in conditional form (as opposed to the bilateral reduction sentences) are true only for a limited domain (e.g. all gases, all light-rays). In the context of an explanation, therefore, we would have a general law very much resembling the ones employed in (I), (II), and (Ib), e.g. $(x)\ [R_x \Rightarrow D(a=kb)_x]$, where R: 'x is a light-ray'.

103Indeed, the case of a law in the form of an equation was not discussed by Carnap in this context.

104Carnap [1936/37], p. 450.

105Carnap's example (very much akin to ours) consists of the three reduction sentences introduced in connection with the different modes of measuring the intensity of an electric current (which were cited above on p. 105).

106Carnap [1936/37], pp. 445f.

107Carnap [1956], p. 64. - Carnap does not sufficiently appreciate this in *Testability and Meaning*. Indeed, it seems that only the criticisms of Pap [1963], pp. 575ff., Hempel [1963], p. 689, and Bridgman [1938] drew his attention to this point. Hempel's argument makes it clear that the problem arises not only from sets of bilateral reduction sentences, but also from an ordinary (non-bilateral) reduction pair as defined above in note 49. Indeed, the problem can be avoided only if a given term has only one bilateral reduction sentence, i.e. if it is a 'pure disposition term' in the sense of Carnap's [1956].

108Thus, if we want to maintain that the concepts introduced by *n* reduction sentences are coextensive and designate just one property, we are required to conjecturally propose all $n(n-1)$ hypotheses in the set (H:). -- Bridgman [1938], p. 122 recommended the adoption of this approach:

> The equivalence of two operations is established by experiment and we must always adopt the attitude that the results of such an experimental proof may be subject to revision [...]

109After all, $(x)\ (S_{1x} \vee S_{2x} \vee S_{3x})$ is not an empirically true proposition.

110These conditionals are entailed by the respective bilateral reduction sentences.

[111] I hope to clarify this contention in the ensueing discussion. Compare Carnap [1956], p. 48:

> the specification, not only of the rules *C* [correspondence rules], but also of the postulates *T*, is essential for the problem of meaningfulness. The definition of meaningfulness must be relative to a theory *T*, because the same term may be meaningful with respect to one theory but meaningless with respect to another.

I suggest that one can read this as "the same term may fulfill condition (D*) with respect to one theory and not with respect to another". To be sure, Carnap seems to have envisioned a more limited role for meaning postulates:

> the theory *T* which is here presupposed in the examination of the significance of a term, contains only the postulates, that is, the fundamental laws of science, and not other scientifically asserted sentences, e.g., those describing single facts. (Carnap [1956], p. 51)

[112] I am disregarding here that, in the present case we happen to have two definitions for NE: since E and NE form an exclusive and exhaustive set of all solubles, the second definition in the present case is (NE <=> -E). The scenario envisioned here would lead to an attribution by definition of E. Therefore, we would end up with conflicting attributions of NE and -NE anyway. -- It is not important, though, whether this sort of case is an exception or the rule - since it will be shown that the the choice of definition vs. set of reduction sentences is only *symptomatic* as far as the significant underlying difference between (i) and (ii) is concerned. We are not dealing with intrinsic properties of definitions versus bilateral reduction sentences.

[113] It should be noted that to treat the conditional not as a meaning postulate but as a falsifiable empirical hypothesis would put all of our theory in serious jeopardy. Just consider: if we remove that meaning postulate (or hypothesis), the definition of NE would still be unchanged. Thus, if we should consequently encounter a state of affairs in which

$[(S_{1a} \& -R_{1a}) \& (S_{2a} \& R_{2a})]$ obtains,

we would still have to attribute NE even though an empirical connection between solubility at low temperatures and having property NE is nowhere subject of empirical law. Furthermore, it would require us to also remove the meaning postulate (or hypothesis) stating that substances with property NE form solutions which do not conduct electricity well. And so forth. (One might argue that this scenario seems too outrageous, since scientists in a situation like this would certainly modify their definition. This is probably true. However, once one is ready to give up the definition, one in effect reverts to a set of bilateral reduction sentences, i.e. to option (i). The scenario discussed in this note, though it may be unlikely, is framed for option (ii).)

114 Carnap [1956], p. 66.

115 Carnap [1956], pp. 68f.

116 Carnap [1956], p. 67.

117 Carnap [1956], p. 69.

118 The distinction is clear enough. I am not sure, however, whether it sustains the contrast between 'observational' and 'theoretical' terms. After all, simple observational terms are particularly well entrenched and characterized by a very great number of reduction sentences - and would thus have to be considered disposition terms which have turned into theoretical terms. Thus, Popper [1963], p. 211 suggests that the term 'water' "is dispositional in perhaps even a higer degree" than 'soluble', where the "higher degree" is due to the greater number of reduction sentences which makes 'water' more theoretical than 'soluble'. Maybe the 'observational' - 'theoretical' juxtaposition can be restored by stipulating that all 'non-problematic' theoretical terms (e.g. all those for which hypotheses (H:) in (i) are true) function as observational terms. -- Concerning the distinction of pure dispositions and theoretical terms, compare also note 111 above and Carnap [1963], p. 950:

> Today I would think [...] that the question of whether a given disposition is *in itself* a disposition, has no clear meaning. I would relativize the term 'disposition' with respect to a language, as Pap suggests. Disposition terms of a given language are then characterized by the fact that they are introduced into this language in a certain way.

Considering all of this together, it becomes apparent that Carnap [1956] with its emphasis on theoretical terms represents not so much a change of opinion but rather a change of perspective in respect to Carnap [1936/37] with its emphasis on (pure) disposition terms: while the earlier article focusses on the introduction of terms, the later one on how to characterize terms within a fully developed theoretical structure.

[119]See above, note 108.

[120]Maybe it should be pointed out here, that this is not a criticism of the notion of independent testability. There is nothing like a vicious circle here. I am simply trying to show that it takes a negotiated consensus in order to consider a law independently testable - and that this negotiation has to go beyond the contemplation of (Ib) and the number of reduction sentences for NE.

[121]This is not to say that analogous hypotheses (H:) were *shown* to be true for Lavoisier's brand of chemistry. -- On the tenuous process of negotiation across (incommensurable) paradigms in the case of phlogistic and 'modern' chemistry, see Nordmann [1986].

[122]I believe that from the point of view of the history of science, it is a major result of the approach adopted here that one can call phlogistic chemistry 'scientific' even though its key concept was ultimately proven to be empirically insignificant. For further examples of non-referential theories see Laudan [1984], pp. 103ff. and especially p. 121. -- That independent testability and condition (D*) have to be relativized to the state of negotiation over empirical hypotheses (H:) at any given time is undesirable only in respect to a quest for a 'logical grammar' of scientific predicates, i.e. for a theory of scientific language which tries to avoid being contingent upon the results of scientific inquiry. As to the 'desirability' of our results, we are now in a situation very similar to the one concerning the proposal for dealing with the tacit immutability requirement: Pap [1963], p. 570 complains in respect to the immutability requirement:

> But even if solubility were a permanent disposition which is never acquired nor lost by a substance, this would be a contingent fact, not something implicit in the 'logical grammar' of disposition predicates.

I am here trying to turn this argument against Pap by posing the rhetorical question: "What is there for the philosophical reconstruction of scientific language to accomplish, if it does not elucidate the forms in which results of scientific inquiry into contingent matters of fact are stated at various stages of theory-development?".

[123]I am suggesting here that 'universal gravitation' was not an independently testable predicate in the sense of (D*). Whether Newton's *hypotheses non fingo* serves as evidence for this suggestion hinges upon issues such as i) whether 'independent testability' presupposes

entrenchment in such a manner that an instantion of the entrenched predicate appears in the explanandum of a higher-level explanation, i.e. that it does not only serve explanatory purposes but is also itself explained; or ii) whether the scope of the original explanandum (which in turn generated abductive awe and thus partially motivated initial acceptance) turns all that is contained in the explanandum into 'old evidence' which cannot serve a significant confirmational function. For the case of 'universal gravitation', I cannot address these issues here. The ensueing chapter will contain a closer analysis of a somewhat related episode in the history of science.

124 In most contexts the term 'diposition predicate' is used for predicates to which a problem-researching attitude has been adopted, i.e. for disposition terms which are presently considered defective in view of the research goal which has not been reached yet. See Levi [1978], pp. 318f.:

> In many contexts, disposition terms - especially explicit disposition terms - do not meet the standards for terms to be used in in fundamental explanation adopted at the time. But investigators may, nonetheless, rest content with such explanations for the moment recognizing the need for further investigation. It is in this sense alone that disposition predicates are to be construed as placeholders.

The placeholder-view of disposition predicates has been argued most forcefully by Levi and Morgenbesser [1978], p. 402, and Levi [1980], pp. 237ff. Similar statements can be found in Quine [1978], pp. 156f., Armstrong [1978], pp. 419ff., Tuomela [1978], pp. 427 and 431, Mellor [1978], pp. 67f., Squires [1968], p. 47, Stevenson [1969], p. 198, and Ringen [1982], p. 123.

125 This juxtaposition of 'placeholder-term' and 'stopgap-device' is a bit unorthodox. Indeed, the two terms are often used interchangeably. I have chosen this juxtaposition for lack of a better idea (the German would allow the juxtaposition of *Platzhalter-* and *Leerstellenbegriff*).

126 At this point, it should be pointed out again that this theory of alternatives in science is inspired by a brief passage in Levi [1967], p. 194:

> whether or not a predicate is to count as dispositional depends in part on theoretical commitments. Thus whether or not motive-attributions are dispositional depends upon whether an investigator considers explanations

that contain motive-predicates to be in need of further explanation in terms of some more 'fundamental' theory (e.g., a physiological theory). Given this [...] account of dispositionality, three distinct kinds of disposition predicates can be listed:

(i) *Problem-raising disposition predicates*: No basis has yet been specified, but the need for such a basis is recognized.

(ii) *Mystery-raising disposition predicates*: No basis is specified, no basis is held necessary, but the predicate is still considered dispositional. [...] (as in the case of 'has dormitive power').

(iii) *Non-problematic disposition predicates*: A basis has been specified. In this case, the disposition predicate can be taken as an abbreviation of an adequate description of the basis.

The account offered in the present chapter differs from Levi's presentation only in the subdivision of (i) into problem-research and problem-deferring predicates (where the inclusion of the latter emphasizes that even pure disposition-terms can function as causally explanatory).

[127]Accordingly, assessments concerning the progressiveness or degeneracy of a research programme cannot be made relative to the present location of a predicate in the chart, but only by considering that location together with the scientists' own expectations, rivaling programmes, etc.

[128]Again, we are reminded of Peirce's perceptive discussion of 'deduction' as a tenuous and difficult operation on a newly proposed general law in conjunction with appropriate background knowlegde, rather than as a simple (atemporal) recognition of the consequences entailed in what is already known (as quoted above on p. 64)

The Case of 'Natural Selection'

The preceding account of scientific activity was designed to serve as a very broad representational scheme which is adaptable to scientific negotiation at any stage of theory-development. It thus predicts only in a very weak, near-trivial sense that any episode in the history of science can indeed be represented within the theory of alternatives in science. Therefore, it is not the purpose of the following venture into the history of evolutionary biology to somehow confirm that account. But the inquiry into the negotiations concerning Darwin's theory of natural selection may exemplify the classificatory power or taxonomic usefulness to the scientist of science of the theory developed in the preceding chapter. Since the ensueing discussion of that episode in the history of biology also prepares that case-study for a probe (in the concluding chapter) into decision-making, argumentation, and rationality in science, the *heuristic* usefulness of that conceptual framework for the business of rational reconstruction may also become apparent. And finally, there is some additional fall-out along the way as conceptual issues concerning the theoretical status of Darwin's theory can be clarified.

A review of neither the primary nor the secondary literature on the case of Darwin and 'natural selection' is possible here. Instead, only the most important stages of the argument and the historical development of Darwin's theory shall be highlighted[1].

i.

In 1859 Charles Darwin published his *On the Origin of Species by Means of Natural Selection*. Rarely has the title of a book so aptly (and plainly) captured its content. For indeed, the book is (in Darwin's own words) "one long argument from beginning to end"[2] which aims to establish that 'species' originated *by means of* natural selection. At times, that argument appears somewhat baroque[3], but nevertheless it is just one, clearly focussed argument. Its aim is to establish an *explanation* of the origin of species by means of natural selection. Now, the very notion of a natural origin implies the mutability of species. It is therefore generally agreed that the evolution of species (including their speciation and adaption) should be considered the explanandum of Darwin's argument, along with e.g. geographical distribution. Darwin's argument thus provides an explanatory mechanism for "large classes of facts" which for the most part he did not have to discover himself:

> No educated person, not even the most ignorant, could suppose that I mean to arrogate myself the origination of thee doctrine that species had not been independently created. The only novelty in my work is the attempt to explain *how* species became modified, & to a certain extent how the theory of descent explains certain large classes of facts; & in these respects I received no assistance from my predecessors.[4]

The limited claim to originality precisely delimits the scope of Darwin's relevant contribution: the achievement of the *Origin* lay in the introduction and defence of 'natural selection'. Accordingly various reconstructions of Darwin's argument have shown the deductive relation between the explanatory theory of natural selection and the (phenomenological) theory of evolution which is part of the explanandum[5]. Michael Ghiselin, in particular, has emphasized the deductive-nomological structure of the argument:

> The theory supposes that there are variations and this is an obvious fact. It also states that for a given structure, sometimes, two different variants have different effects in furthering the lives of the individuals. Thus it is predicted that *if* there are variations, *if* these are inherited, *if* one variant is more suited to some task than another, and *if* the success in accomplishing that task affects the ability of the organisms to survive in whatever happens to be their environment, *then* natural selection will produce an evolutionary change. Such conditional statements are basic to the Darwinian theory, and this conditionality has given rise to much confusion.[6]

One source of the confusion is the lack of a clear referent for the term 'natural selection' in this statement. While it appears in the consequent of Ghiselin's long conditional, it also designates some further condition which in conjunction with the clauses of the antecedent produces evolutionary change. On the account adopted in the previous chapter, a somewhat more intuitive deductive-nomological reconstruction of Darwin's argument recommends itself. It construes 'natural selection' as a disposition predicate, that is, as a predicate introduced by a reduction sentence which (like Ghiselin's formulation) assigns the term 'natural selection' to a conditional consequent of a conditional. This construal of 'natural selection' as a disposition predicate takes its clue from the following passage of the *Origin* and it is borne out (at least, not contradicted) by the entire work and its subsequent development in scientific negotiation.

> Several writers have misapprehended or objected to the term Natural Selection. Some have even imagined that natural selection induces variability, whereas it implies only the preservation of such variations as occur and are beneficial to the being under its conditions of life. No one objects to agriculturalists speaking of the potent effects of man's selection; and in this case the individual differences given by nature, which man for some object selects, must of necessity first occur. Others have objected that the term selection implies conscious choice in the animals which become modified; and it has even been urged that as plants have no volition, natural selection is not applicable to them! In the literatal sense of the word, no doubt, natural selection is a false term; but who ever objected to chemists speaking of the elective affinities of the various elements? - and yet an acid cannot strictly be said to elect the base with which it will in preference combine. It has been said that

> I speak of natural selection as an active power of Deity; but who objects to an author speaking of the attraction of gravity as ruling the movements of the planets? Every one knows what is meant and is implied by such metaphorical expressions; and they are almost necessary for brevity. So again it is difficult to avoid personifying the word Nature; but I mean by Nature, only the aggregate action and product of many natural laws, and by laws the the sequence of events as ascertained by us. With a little familiarity such superficial objections will be forgotten.[7]

In the *Origin*, Darwin introduced the disposition predicate 'natural selection' and he attributed it to 'Nature'. In full, that disposition attribution may read somewhat like this:

Nature is disposed to select the fittest.

However, at the time when Darwin was writing the *Origin*, he provided an even more extensive and still more accurate interpretation of 'natural selection' in dispositional terms.

> I had not thought of your objection of my using the term "natural selection" as an agent. I use it much as a geologist does the word denudation – for an agent, expressing the result of several combined actions. I will take care to explain, not merely by inference, what I mean by the term; for I must use it, otherwise I should incessantly have to expand it into some such (here miserably expressed) formula as the following: "The tendency to the preservation (owing to the severe struggle for life to which all organic beings at some time or generation are exposed) of any, the slightest, variation in any part, which is of the slightest use or favourable to the life of the individual which has thus varied; together with the tendency to its inheritance."[8]

It is not only the word "tendency" which supports the dispositional construal of 'natural selection'. From this passage also emerges a structural feature shared by this and other disposition predicates. Darwin speaks of 'natural selection' as "the result of several combined actions". At the same time, however, he treats this 'result' as a cause:

> Natural Selection almost inevitably causes much Extinction of the less improved forms of life, and induces what I have called Divergence of Character.[9]

While further research will have to elaborate the precise character of the compositional effect, for now, reference to that effect can serve explanatory purposes.

Like all disposition predicates, 'natural selection' is introduced by a reduction sentence, i.e. together with a corresponding law that is true of all those states which have the disposition to '(naturally) select the fittest':

$$(x)(t)\ (S_{xt} \Rightarrow [NatSel_x \Leftrightarrow R_{xt}])$$[10]

'S_x' stands here for a conjunction of two stimulus-conditions: variation and heredity. 'R_x' encompasses a whole range of partly or entirely overlapping response-phenomena: evolution, speciation, extinction, adaptation, geographical distribution, reversion, and variability[11]. Thus, if there is variation and heredity, then the property to select the fittest obtains if and only if there is evolution, adaptation etc.[12] So introduced, 'natural selection' can fulfill its explanatory purpose. To say that a species underwent an evolutionary change *because* of natural selection is tantamount to proposing an explanatory scheme of type (I) (see above, p. 89), where 'E' is an environmental state of 'nature':

$$(x)\ (E_x \Rightarrow NatSel_x)$$
$$(x)\ (S_x \Rightarrow [NatSel_x \Leftrightarrow R_x])$$
$$E_a$$
$$\underline{S_a \qquad\qquad}$$
$$R_a$$

That is, a species underwent an evolutionary change *because* the following conjunction of circumstances hold: every environmental state of nature has the disposition to select the fittest; having that property means for a given state (at least) that *if* variation and heredity obtain in that environment, *then* a species will undergo an evolutionary change; and there is environment *a* and variation and heredity obtain in it. To be sure, the explanation as stated is deficient somewhat like Ghiselin's conditional. The role and meaning of 'natural selection' in this explanation is not clear. Indeed, all that one can say about 'natural selection' is that it designates a site for further research, namely the site that is only vaguely described as an environmental state in which variation and heredity obtain[13]. After all, Darwin did not put forth the unqualified conditional

$$(x)\ (E_x \Rightarrow [S \Rightarrow R])$$

which is entailed by the general law and the reduction sentence of his explanandum but which does not capture that crucial perspective towards further research[14]. Darwin's formulation calls for an empirical clarification of what it means for an environmental state to have the property to 'naturally select the fittest'. This pursuit will be satisfied (as in the case of sugar and solubility) once one is able to say what *else* must be true of an environmental state that has that property, that is, what else besides 'variation', 'heredity', 'evolution', 'adaptation' etc. as these terms appear in the reduction sentence for 'natural selection'. In short, 'natural selection' appears as a placeholder or stopgap device in the explanatory hypothesis proposed by Darwin. An empirical inquiry of that property will adopt a placeholder view of 'natural selection' and it will aim for independent testability.

ii.

Undoubtedly, the first question to be asked about this reconstruction of Darwin's argument concerns its advantages compared to other reconstructions. Once again, for present purposes the criterion of comparison lies in the historiographic power or utility of the competing reconstructions. And yet, in this case the historiographical virtues happily coincide with a conceptual clarification as one looks at the issues which Darwin had to address after introducing 'natural selection'.

Just like the dispositional property 'soluble', 'natural selection' is either a placeholder or stopgap device in Darwin's scientific explanation. If one returns to the passage from the *Origin* which suggested the dispositional interpretation of 'natural selection', a further analogy between 'solubility' of 'dormitive power' and 'natural selection' emerges. In it, Darwin complains that

> It has been said that I speak of natural selection as an active power of Deity; but who objects to an author speaking of the attraction of gravity as ruling the movements of the planets? Every one knows what is meant and is implied by such metaphorical expressions; and they are almost necessary for brevity.[15]

The complaint is directed against a mystery-raising interpretation of his theory. To fend off teleological readings of any sort was one of the difficult tasks before him. As in the case of 'solubility' and 'dormitive power' the outcome of the choice between the problem-raising and the mystery-raising attitude it not determined one way or the other by the formulation of the explanatory account itself. Indeed, questions as to whether evolutionary biologists fall back into a mystery-raising attitude have been raised intermittently ever since the inception of Darwin's theory[16].

While modern philosophers have largely overcome the worry of Darwin's contemporaries that the theory of natural selection might be teleological[17], they have introduced the specter of another mystery-raising interpretation of Darwin, one that has also been raised in connection with 'solubility', but one that has hardly worried Darwin and his contemporaries[18]: they have suggested that the theory as presented by Darwin is (in crucial aspects, at least) tautological, e.g. that 'adaptation' defines 'fitness' while it is supposed to explain 'fitness'[19]. Now if one looks at the reduction sentence (a conditional definition) for 'natural selection', the plausibility as well as severe limitations of this interpretation become quite apparent. The presumed tautological character of the theory arises as soon as one drops the antecedents or test-conditions of the reduction sentence. The temptation to do so is small in the case of 'solubility'. It looms large in the case of 'natural selection' since - as Ghiselin said - "there are variations [and heredity] and this is an obvious fact". However, if the antecedent in the reduction sentence for 'natural selection' is always satisfied, that is quite contingent and an empirical matter. Accordingly one truncates the reductions sentence at one's own risk, especially since the history of evolutionary biology has shown that the clarification of 'natural selection' is closely linked to a better understanding of the processes involved in variation and heredity. If one takes into account the whole reduction sentence and the corresponding empirical laws, there is clearly nothing tautological about the theory of natural selection. But to be sure, once variation and heredity are taken for granted, Darwin's placeholder predicate or stopgap device does not provide a particularly informative internal relation between e.g. fitness and adaptation. Thus, to claim that the theory is tautological is wrong even if one limits it to Darwin's original formulation. On the other hand, the predicate 'natural selection' should not be taken to denote anything deeper than a well-delineated area for future research[20].

The problems posed by 'natural selection' can thus be framed in analogy to the problems posed by 'solubility'. And incidentally, doing so helps to clarify a few persisting conceptual problems.[21]

iii.

In the passage which was cited above, Darwin chose the label "metaphorical expression" for "natural selection". It proves to be a very useful label in a context of negotiation where the introduction of a disposition predicate by one sweeping reduction sentence is at stake. For, it firstly captures its openness for interpretation and the need for its clarification. That is, it emphasizes that one has to adopt a problem-raising attitude towards it. For, unquestioning acceptance of that predicate for explanatory purposes would amount to a mystery-raising treatment. By the same token, the label "metaphoric expression" remains uncommitted as far as the next alternative is concerned, i.e. the question of whether to adopt a problem-researching or a problem-deferring attitude. Darwin's position on this matter is particularly clear in a letter to Hugh Falconer:

> You make important remarks *versus* Natural Selection, and you will perhaps be surprised that I do to a large extent agree with you. I could show you many passages, written as strongly as I could in the *Origin*, declaring that Natural Selection can do nothing without previous variablility; and I have tried to put equally strong that variability is governed by many laws, mostly quite unknown. My title deceives people, and I wish I had made it rather different. [...] – for years I was stopped dead by my utter incapability of seeing how every part of each creature (a woodpecker or swallow, for instance) had become adapted to its conditions of life. This seemed to me, and does till seem, the problem to solve; and I think Natural Selection solves it, as artificial selection solves the adaption of

> domestic races for man's use. But I suspect that you mean something further, - that there is some unknown law of evolution by which species necessarily change; and if this be so, I cannot agree. [...] If, indeed, an elephant could succeed better by feeding on some new kinds of food, then any variation of anykind in the teeth which favoured the grinding power would be preserved. Now, I can fancy you holding up your hands and crying out what bosh! To return to your concluding sentence: far from being surprised, I look at it as absolutely certain that very much of the Origin will be proved rubbish; but I expect and hope that the framework will stand.[22]

Here, a rejection of a mystery-raising interpretation of 'natural selection' is repudiated[23], while a placeholder-attitude towards it is forcefully advocated. In due course, the framework will be filled out, and future research will have clarified the meaning of 'natural selection':

> As in time the term must grow intelligible the objections to its use will grow weaker and weaker.[24]

But Darwin also argued that 'natural selection' may be interpreted as a stopgap-device. To this end he frequently employed references to the history of physics:

> I believe that this view in the main is correct, because so many phenomena can be thus grouped together and explained. But it is generally of no use; I cannot make persons see this. I generally throw in their teeth the universally admitted theory of the undulation of light, - neither the undulation nor the very existence of ether being proved, yet admitted because the view explains so much.[25]

Darwin's argument for the tentative acceptance of 'natural selection' in the absence of direct evidence has to remain somewhat indifferent to the choice between the problem-researching and the problem-deferring attitudes. Indeed, to enable that latter exchange, 'natural selection' had to be made palatable to the minds of those who would want to make it subject of immediate investigation as well as to those who would see in it a cogent explantory theory.

iv.

So much for Darwin's contribution to the negotiation of 'natural selection'. The rest is history - and the overall pattern of its further development is also representable within the scheme developed in the previous chapter.

The negotiation concerning a preferred interpretation for the disposition predicate 'natural selection' continued well into this century. With the temporary demise of Darwinism[26] evolved a controversy on the choice between placeholder and stopgap interpretation. The sturdiest defenders of Darwinism had adopted a Humean notion of causality in terms of constant conjunction. On this notion of causality, adoption of a stopgap, problem-deferring attitude towards 'natural selection' amounted to waiving all speculation concerning the processes underlying natural selection. Instead, the efficacy of natural selection was measured with statistical methods. In accordance with the idea of 'constant conjunction', the efficacy of natural selection is presumed to consist of continuous small-case variation which becomes perceptible as it cumulates. In the words of W.F.R. Weldon, one of the main proponents of this 'biometrical' point of view,

> numerical data [...] contain all the information necessary for a knowledge of the direction and rate of evolution. Knowing that a given deviation from the mean character is associated with a greater or less percentage death-rate in the animals possessing it, the importance of such a deviation can be estimated without the necessity of inquiring how that decrease or increase in the death-rate is brought about, so that all ideas of 'functional adaptation' become unnecessary.[27]

The biometricians' statistical and purely phenomenological rather than experimental methods encountered harsh criticism in the scientific community: The biometricians'

> methods of attempting to penetrate the obscurity which veils the interactions of the immensely complex bundle of phenomena which we call crab and its environment, appear to me not merely inadequate, but in so far as they involve perversion of the meaning of accepted terms and a deliberate rejection of the method of inquiry by hypothesis and verification, injurious to the progress of knowledge.[28]

The opponents of the biometricians wanted to gain an understanding of the causal processes underlying or determining natural selection, they held a placeholder view of 'natural selection'. With the rediscovery of Mendel's work in genetics, these opponents of biometrics had found their home. Historians of the controversy therefore juxtapose the parties to the conflict in the following terms:

> In Biometry, the extent of hereditary resemblance, and thus the strength of heredity, was given directly, *by definition*, as the degree of manifest resemblance between "phenotypes" (as we would say). For a Mendelian, on the other hand, such superficial resemblance could only be of significance as an indicator of the operation of Mendelian factors. The latter were the source of all inherited resemblance or variation. Only explanations in terms of these factors ultimately counted as contributions to knowledge about heredity.[29]

Having found a physically localizable 'cause' for heredity, variation, and evolution, the Mendelians abandoned Darwin's notion that change is continuous and proceeds in small steps. And 'natural selection' became an obsolete notion. But while the Biometricians appeared overly cautious in their avoidance of empirical speculation, the Mendelians overrated the explanatory force of Mendelism. As Julian Huxley observed:

> Bateson did not hesitate to draw the most devastating conclusions from his reading of the Mendelian facts. [...] he concluded that the whole of evolution is merely an unpacking. The hypothetical ancestral amoeba contained - actually and not just potentially - the entire complex of life's hereditary factors. The jettisoning of different portions of this complex released the potentialities of this, that, and the other group and form of life. Selection and adaptation were relegated to an unconsidered background. [...] The biometricians stuck to hypothetical modes of inheritance and genetic variations on which to exercise their mathematical skill; the Mendelians refused to acknowledge that continuous variation

> could be genetic, or at any rate dependent on genes, or that a mathematical theory of selection could be of any real service to the evolutionary biologist.[30]

Huxley presents his view of the controversy under the heading "The Eclipse of Darwinism". Statistical research upon a stopgap interpretation of 'natural selection' was degenerating and threatening to become a self-serving and mystery-raising mathematical exercise. The genetical research programme which was to render the placeholder predicate 'natural selection' obsolete, appeared to take a mystery-raising twist of its own. It was high time for 'the great synthesis' of population genetics which would combine some of the biometricians's methods as well as Mendelian theories[31].

The great synthesis inaugurated a research programme which is still active today. At its foundation lies a placeholder view of 'natural selection'. While 'natural selection' is now an independently testable, theoretical term[32], one is still far from wondering whether molecular and population genetics have rendered it unproblematic[33]. Indeed, some of the research which has led to the independent testability of 'natural selection' has also given rise to new controversy. It will be remembered that the independent testability of theories depends on the acceptance of certain empirical propositions which ensure that in the various contexts one is actually testing for the same property[34]. The contemporary controversy on the 'units' and 'levels' of selection exemplifies the difficulty of this negotiation: the move to the genetic level allowed for a radically improved conception of 'natural selection'; at the same time, it led to notions such as the 'selfish gene' which may or may not work towards the preservation and survival of the organism of which it is a part; and scientists are now faced with the question of how to best preserve a unified conception of 'natural selection'.

V.

Once the territory has been staked out and the historical development framed as a succession of choices between certain alternative interpretations of a dispositional concept, the reasoning in defence of these choices can be approached. Given the vastness of material concerning the history of evolutionary biology, this discussion shall be limited to that phase of the negotiation where it was first proposed that 'natural selection' as a problem-raising disposition predicate is worthy of scientific consideration and of admittance into the body of (tentatively) accepted scientific knowledge.

Notes to Chapter 3

[1] The following works proved to be most useful for an understanding of Darwin's argument and the events following upon the publication of the *Origin of Species*. A general introduction is provided by Ruse [1981]. Introducing a fascinating collection of reviews by Darwin's contemporaries, Hull [1983] presents the most concise formulation of the issues. A wealth of material can be found in Ellegard [1958]. Very useful were Ghiselin [1984], Himmelfarb [1959]. -- For Darwin's intellectual development leading up to the publication of the *Origin* such diverse works as Manier [1978] and Gruber [1981] were drawn upon. Of Darwin's own writing, his letters proved to be most important (Darwin [1887] and [1903]) aside from, of course, the *Origin* itself (Darwin [1959]). The primary source for contemporary criticism of the theory of natural selection was Hull [1983].

[2] Darwin [1887/1893], Vols. 1, pp. 103/82. [Since, as of yet, there is no standardized way of quoting from the *Life and Letters of Charles Darwin*, all references are to both, the 1887 edition in 3 volumes and the 1893 edition in 2 volumes. When a letter is quoted, date and addressee will also be stated. Since the passage just quoted comes from Darwin's autobiography, see also Darwin [1958], p. 140.]

[3] See e.g. the disarming confessions concerning the unintelligibility of certain parts of the argument in Gale [1982]. Compare also Huxley [1906], p. 301.

[4] Darwin to Baden Powell in 1860, as quoted in Young [1971], p. 445.

[5] See especially Ruse [1971] which closely adheres to the formulations given in the *Origin*. - For a very perceptive (though very critical) reconstruction of the deductive-nomological structure in Darwin's argument by one of Darwin's contemporaries see Hopkins [1983]. Hopkins introduces the distinction between phenomenological laws (like Kepler's laws of planetary motion) and physical laws (like Newton's theory of universal gravitation). While a theory of evolution by itself would be phenomenological, Hopkins considers Darwin's theory of natural selection a physical law which properly identifies causes. To be sure, he deems it a bad physical law and thus concludes on p. 272:

> Biological science requires at present its Keplers rather than its Newtons - the discovery

of the more obvious laws according to which its phenomena may be arranged, rather than attempts at that higher generalization which may account for such laws by the operation of physical causes.

[6] Ghiselin [1984], pp. 64f.

[7] Darwin [1959/1962], chapter IV, 14/pp. 91f. [Again, as there is no standard format for references to the *Origin*, various editions of varied accessibility and practicality will be quoted. If (as in the present case) the quoted passage did not yet appear in the first edition of the *Origin* there will be a reference to the chapter and sentence number in Peckham's 1959 *Variorum* edition and to the 6th edition as it appeared in paperback in 1962. If a passage appeared in the first edition already, a third and fourth reference will be added, one to the page number of the first edition of 1859 and another to a popular reprint of that first edition in paperback in 1981.] - This brief passage is so rich that one may well say that the remainder of this chapter is devoted to its interpretation, and *a fortiori* that countless scientists during the last 125 years have negotiated its meaning.

[8] Darwin [1903], letter # 79 to Asa Gray (November 29th, 1859), pp. 126f. -- Another very explicit discussion of the dispositional character of 'natural selection' can be found in an exchange between Owen [1868] and Romanes [1896]. Owen [1868], p. 794 considers all dispositional explanations on a par with the 'dormitive power' explanation which was parodized by Moliere:

> Natural Selection is an explanation of the process [of transmutation] of the same kind and value as that which has been proffered of the mystery of "secretion." For example, a particular mass of matter in a living animal takes certain elements out of the blood, and rejects them as "bile." Attributes were given to the liver which can only be predicated of the whole animal; the "appetency" of the liver, it was said, was for the elements of bile, and "biliosity," or the "hepatic sensation," guided the gland to their secretion. Such figurative language, I need not say, explains absolutely nothing of the nature of bilification.

Romanes [1896], p. 334 responds to this passage by stressing the problem-raising attitude which can be chosen towards 'natural selection' and which sets it apart from the explanation cited by Owen:

> it was little less than puerile in him [Owen] to see no more in the theory of natural selection than such a mere figure of speech. To say that the liver selects the elements of bile, or that nature selects specific types, may both be equally unmeaning re-statements of facts; but when it is explained that the term natural selection, unlike that of "hepatic sensation," is used as a shorthand expression for a whole group of well-known natural causes - struggle, variation, survival, heredity, - then it becomes evidence for an almost childish want of thought to affirm that the expression is figurative and nothing more.

Though Romanes's retort is also not as clear as it might be, the repeated use of the word "more" is particularly noteworthy: Romanes does not repudiate Owen's analogy, but with the adoption of the problem-raising attitude there is *more* to 'natural selection' than Owen appears to see.

[9] Darwin [1959/1962/1859/1981], Introduction, 39/p. 28/p. 5/p. 68.

[10] The corresponding law over the domain of all those things which have the disposition 'natural selection' is, of course, $(x)(t)\ (S_{xt} \Rightarrow R_{xt})$.

[11] Here, 'variability' designates the limited scope of variation and not the ability to vary. Darwin referred to the latter in the passage just quoted. Variability and reversion are what Popper calls the "conservative principles" of evolution (see Popper [1976], p. 170).

[12] For the distinction between 'fitness' (as belonging to the explanans) and 'adaptation' (belonging to the explanandum), see Campbell [1983]. A remark on p. 64 illuminates the notion that 'selection of the fittest' is a disposition of nature:

> Darwin's theory *postulates* that individuals vary in their fitness value; it *explains* why individual characteristics appear to be so well adapted to the environment.

[13] Compare note 76 to chapter 2 on pp. 144f. - Of course, Darwin went a little bit further than this by suggesting possible selective mechanisms for which he uses the generic label 'struggle for existence'. His largely phenomenological and anecdotal exploration of that 'struggle for existence' was a first contribution to the research pogramme which he inaugurated by proposing his theory of natural selection.

[14] Introducing an interesting twist to the familiar charge that Darwin's theory is tautological, Popper suggests that this statement is true as a matter of 'situational logic' (Popper [1976], p. 168), i.e. that the notion of 'natural selection' is redundant. However, as Popper points out, not every possible situation E will in the terms of situational logic entail the truth of (S => R). It just so happens that the situations called 'nature' or 'physical environment of biological organisms on earth' *do* entail the truth of that conditional. Then, however, one has to pose empirical questions concerning what it is about these situations that makes them situationally entail (S => R) - and one will soon rediscover 'natural selection' as the property shared by these situations.

[15] Darwin [1959/1962], chapter IV, 14/p. 92.

[16] And nowadays they have to be raised most vehemently against the sociobiological enterprise (see Lewontin et.al. [1984]). For a lucid discussion of the difficulties associated with the investigation of 'goal-directed processes' as well as 'functional explanations' in biology see Nagel [1979b], pp. 275-341. He maintains that these notions do not necessarily involve a mystery-raising attitude or untenable teleological presuppositions. Other philosophical contributions to the ongoing negotiation on how to combine functional explanation with a problem-raising attitude are Nagel [1979], pp. 398-446, Hempel [1965], pp. 297-330, and Cummins [1975]. And for a philosophical contribution by a preeminent biologist see e.g. Mayr [1976], pp. 383-404.

[17] Few of Darwin's contemporaries addressed this issue. And when they did, they often did not worry about it themselves but suggested that an apparent materialist like Darwin should. Wollaston [1983], pp. 133f. expresses that line of criticism forcefully as he challenges the attribution of a disposition (interpreted as a placeholder or stopgap device at that) to 'nature'.

> We believe it was Coleridge who first called attention to this fact, that to treat a mere abstraction as an efficient cause is simply absurd. [...] But who is this "Nature," we have a right to ask, who has such tremendous power, and to whose efficacy such marvellous performances are ascribed? What are her images and attributes, when dragged from her wordy lurking-place? Is she aught but a pestilent abstraction, like dust cast into our eyes to

obscure the workings of an Intelligent First Cause of all?

On the basis of a largely correct analysis of Darwin's theory, Wollaston adopts a mystery-raising attitude towards it and rejects it as such - since his need for mysteries is already satisfied by an "Intelligent First Cause". Compare also: Hull [1983], p. 153 and Hopkins [1983], p. 271. The issue was also raised in a letter by Sedgwick to Darwin (December 1859) in Darwin [1887/1893], Vols. 2, pp. 247-250/42-45.

18 One of the few remarks by Darwin which addresses the charge that his theory is tautological or expresses a truism of situational logic (see above, note 14) also happens to be one of his few cutting remarks. It can be found in a letter to W.H. Harvey (August 1860) in Darwin [1903] Vol. 1, p. 161, letter # 110:

> The upshot of your remarks at p. 11 is that my explanation, etc., and the whole doctrine of Natural Selection, are mere empty words, signifying the "order of nature". As the above-named clearheaded men, who do comprehend my views, all go a certain length with me, and certainly do not think it all moonshine, I should venture to suggest a little further reflection on your part.

19 Of course, a good number of Darwin scholars have mounted a vigorous defence on this score. For instance, Campbell [1983] aptly concludes her discussion (p. 65):

> Adaptations pertain to individual characteristics whose function is often evident. [...] Fitness is a holistic concept. It refers to reproductive wholes. Fitness and survival are not tautological concepts since it is phenotypic individuals who are fit or not, while it is genes that survive.

20 Ghiselin [1984], p. 69 warns quite polemically against overinterpreting 'natural selection':

> The concept of natural selection seems unsatisfactory simply because it is intelligible by reason.

For the present purposes, "intelligible by reason" should be substituted by "a sensible but tentative foundation of a research programme".

21 Indeed, a third contemporary issue in philosophical discussions of Darwin's argument is diffused as one introduces 'natural selection' by a reduction sentence which functions somewhat like an inference ticket in deductive-nomological explanation. The question is

whether Darwin's theory should be interpreted 'syntactically' (as a partly interpreted axiomatic theory) in terms of explanatory power or 'semantically' in terms of semantic adequacy. Thompson [1983] provide a semantic approach towards modern evolutionary biology, while Lloyd [1983] deals with Darwin's theory and his support for it. Though Lloyd provides a splendid reconstruction of Darwin's argument, it is quite misleadingly framed in that syntactic-semantic dichotomy. As Lloyd herself points out (p. 119), there is considerable overlap between the two approaches to science when it comes to writing its history. This overlap reflects the absence of a real issue here. The construal of 'natural selection' as a disposition predicate that is employed in deductive nomological explanation, shows its tremendous explanatory power. Darwin again and again emphasized this aspect of his theory as an argument for its (tentative) acceptance (compare Lloyd [1983], p. 116). However, explanatory power is here exerted by a disposition predicate which is either a placeholder or a stopgap-device. The appeal to 'natural selection' in an explanation with deductive nomological form therefore serves primarily as a heuristic artifact within ongoing scientific deliberation (after all, it is the outcome of abduction). Scientists then deliberate and negotiate on a modified, more elaborate, and meaningful construal of that artifacts. That process of negotiation is best described as a process of fitting theory and empirical data. Lloyd focusses on that process of negotiation:

> When natural selection theory is said to present a set of related models, it is meant that there are certain model types which are given in the theory to account for observed phenomena; the variables of these model types are specified and instantiated through hypothesis and testing in a recursive manner. (p. 118)

It is difficult to see how this process of negotiating and fitting should somehow conflict with intermediary and ultimate formulations of scientific explanations. On this question of philosophical and historiographical compatibilities see also note 9 to chapter 2 on pp. 132f.

22 Letter # 143, October 1st, 1862 in Darwin [1903], Vol.1, pp. 208f.

23 One of his supporters among scientists who came dangerously close to a mystery-raising interpretation was Asa Gray. In his first review (Gray [1963], p. 46) of the *Origin* he "judge[s] it probable" that Darwin

> regards the whole system of Nature as one

> which had received at ist first formation the impress of the will of its Author, foreseeing the varied yet necessary laws of its action throughout the whole of its existence, ordaining when and how each particular of the stupendous plan should be realized in effect, and - with Him to whom to will is to do - in ordaining doing it.

While Darwin does not quarrel with Gray's theological interpretation (compare the motto from Butler which he added to the second edition of the *Origin*), he would not use it as an argument for or against any theory including his own. Gray's attitude appears questionable on this latter score.

24 Letter to A.R. Wallace (# 191), July 5th 1866 in Darwin [1903], Vol. 1, p. 271.

25 Letter to F.W. Hutton (# 124) on April 20th, 1861. In Darwin [1903], Vol. 1, p. 184.

26 See Bowler [1983].

27 Weldon as quoted in Farrall [1975], p. 284.

28 E.R. Lancaster as quoted in Farrall [1975], p. 284.

29 MacKenzie and Barnes [1979], p. 199. Aside from Provine [1971], interpretations of the controversy between Mendelians and Biometricians are provided by Farrall [1975], Roll-Hansen [1980], and Bowler [1983], pp. 38ff. See also Huxley [1942], pp. 22-24 and Punnett [1911], pp. 10ff.

30 Huxley [1942], p. 24.

31 Mayr [1976], p. 346 puts the various controversies into a larger context:

> the genetic argument was merely a symptom of a far deeper disagreement, the choice between saltationism and Darwin's gradual evolution through natural selection. The final reconciliation among evolutionists in the 1930s and 1940s, often designated "the great synthesis," is much more appropriately considered the real end of a controversy that had started well within Darwin's lifetime.

32 In the preface of one of the ground-breaking works towards the "great synthesis", R.A. Fisher criticizes the

practice of conflating the theories of evolution and natural selection. He makes it plain that one of the reasons for holding them apart consists of the then-emerging possibility to subject the theory of natural selection to independent test:

> To treat Natural Selection as an agency based independently on its own foundations is not to minimize its importance in the theory of evolution. [...] In addition it will be of importance for our subject to call attention to several consequences of the principle of Natural Selection which, since they do not consist in the adaptive modification of specific forms, have necessarily escaped attention. (Fisher [1958], p. x, first published in 1929)

On the independent testability of 'natural selection' which became possible only with population genetics see Wassermann [1978]. As should be expected, the introduction of further reduction sentences for 'natural selection' has led to a clarification of its meaning. This is bluntly expressed by Huxley [1942], p. 16:

> The term Natural Selection is thus seen to have two rather different meanings. In a broad sense it covers all cases of differential survival: but from the evolutionary point of view it covers only the differential transmission of inheritable variations.

33 To be sure, there has been discussion whether Darwin's theories are fully expressible within (or reducible to) molecular genetics. Generally, that possibility is assessed with great scepticism (and from a philosophical point of view). See e.g. Beatty [1982].

34 Independent testability, i.e. the satisfaction of condition (D*) depends on the acceptance of the set of hypotheses (H:). See above, pp. 111ff.

Argumentation Repertoires in Scientific Negotiation

A historical test-case has been prepared by placing it within the conceptual framework provided by the theory of alternatives of science. The plan of inquiry which was developed in the first chapter can now be brought to bear on the negotiation concerning 'natural selection' and the business of rational reconstruction can begin.

i.

Scientific negotiation is differentiated from other contexts of decision-making not by some special form of rationality or by the goals which scientists individually or collectively pursue, but by characteristic constraints on the alternatives to be negotiated upon and the arguments that can be adduced in support of a choice among the alternatives. The adopted plan of inquiry accordingly calls for the rational reconstruction of science and the history of science on the grounds of general decision-theory with respect to acceptable arguments. In order to apply (if only informally) general decision theory to the negotiation of the alternatives that were introduced together with 'natural selection', the alternatives themselves have to be made explicit along with the argumentative means available in support of one's choice among them.

As was noted in the previous chapter, Darwin resolved that 'natural selection' should be accepted as a problem-raising disposition-term. Thus, he made a particular choice among clearly defined alternatives, implying that for the time being no decision is needed on whether to adopt a problem-deferring or problem-researching attitude. His choice furthermore implied his unambiguous rejection of a mystery-raising attitude and his readiness to demonstrate that, indeed, he himself does not perpetuate this attitude. At the same time, he could and would not enter into a debate on whether 'natural selection' is still problematic or already unproblematic.

No more needs to be said about the alternatives at hand. What needs to be clarified now is the notion that only arguments drawn from scientifically acceptable argumentation repertoires can be used to defend a choice among the alternatives.

It was pointed out towards the end of the first chapter that knowledge of scientists' goals is irrelevant for an understanding of science. By the same token, the notion of 'rationality' was reduced to 'providing a cogent rationale': it makes no difference what goals scientists have as long as they rationally address scientific alternatives using scientifically acceptable arguments[1]. Thus, a (cogent) rationale adduced for one's choice among alternatives may or may not coincide with one's reasons or motives for making that choice, where repertoires of acceptable arguments constrain only the adduced rationale and not one's reasons or true motives. But regardless of whether reason and rationale do in a given case coincide, employment of arguments from acceptable argumentation repertoires serves to opportunistically embellish a choice which has already been made, if only in a tentative mood or on a hunch. Accordingly, if there is a historical law (embodying a methodological rule) suggested by the discussion so far, it is a law of tempered opportunism in science and it could be formulated like this:

> (L*) (x)(y)(t) (If x is a scientist and y is an option among the set of alternatives in science in respect to a given hypothesis, and x has adopted y, then x will rationally employ [as long as y remains controversial] all arguments which are compatible with the available evidence, which support y, and which can be drawn from a repertoire of acceptable scientific arguments)[2]

The force of this law or rule differs greatly in various phases of theory development. Depending on the powers of human ingenuity, (L*) permits a practically boundless accumulation of arguments in the early phases of theory development. But very few arguments may be both acceptable and compatible with the available evidence in a later phase where e.g. one is testing a hypothesis which has been derived from the theory and appropriate background knowledge and where a rivaling theory has been seriously challenging the theory under investigation. But even in this later phase, imaginative scientists will still be able to provide arguments for the acceptance of that theory, regardless of the experimental outcome. A single admissable argument which - preferably on evidential grounds - warrants or requires a certain choice among alternatives and which accordingly ought to persuade in one fell swoop the entire scientific community is only a limiting case of (L*), and probably a very rare case at that. In contrast, most philosophical theories of science take this limiting case of (L*) as the paradigm case of methodological negotiation[3]. (L*) does not attribute ultimate persuasive to any argument at all. But it attributes importance in the process of persuasion to the *weight* of arguments instead of the quality of a single argument, where 'weight' is a composite of quality and sheer number, the latter reflecting the relative ease with which arguments for a certain choice can be put forth. (L*) can thus explain something that is apparently quite trivial and which many theories of science nonetheless cannot explain, namely the generic phenomenon

(Ph1) Scientists tend to employ more than one argument (or more than one type of argument[4]) in defence of their choice among alternatives in science.[5]

Along with this phenomenon often comes another one.

(Ph2) Short of contradicting themselves, scientists will employ conflicting argumentation strategies.

For instance, scientists critical of Darwin's theory have argued that first of all it is incoherent, nonsensical, tautological, or (empirically) meaningless and that secondly it is contradicted by empirical evidence. And yet these scientists are not engaged in a contradictory interpretation of Darwin's theory since they keep these criticisms separate, trying to anticipate possible responses and arguing, in effect, that e.g. *if* the theory is not tautological, *then* it is false and *vice versa*[6]. The weight of arguments is used to smother the opponent, a single argument will seldom do, even if it is elegantly aims right for the heart of a matter.

To be sure, it is much easier to posit a law like (L*) than to defend it as a confirmable or falsifiable law in the history of science. Indeed, such a defence shall not be attempted here. One of (L*)'s major difficulties lies in its tacit appeal to the epistemic states of scientists at any given time: surely, scientists can draw only on those argumentation repertoires with which they are familiar. Here one encounters the obvious problem of how to find independent grounds for establishing that familiarity. And even if one succeeds in doing that (maybe in reference to ascertainable aspects of a scientist's training or to the certifiable appeal by the same scientist to that repertoire in another context), by far not all is done: one will still have to resolve in any given case to what one should ascribe an apparent 'falsification' of (L*). After all, the failure of making use of a repertoire in a situation where that would be expedient may be due simply to a discrepancy between the historian and the historical agent when it comes to making this assessment of expediency. Similar problems arise when it comes to designating and precisely delimiting the repertoires available at any given time. While there are repertoires which have been publicly certified by philosophers of science, many acceptable arguments stem from repertoires which remain tacit or which are constituted merely by historical precedent. And at the same time, not all philosophically certified methodological contrivances may yield arguments which are straightforwardly acceptable to the scientific community, for instance if they are lacking in intuitive transparency or if they employ controversial presuppositions that enter into their use[7]. These types of problems are closely linked to the general problem of whether historical laws are discernible[8].

Even if a positive answer to these methodological problem id difficult to find, reference to law-like formulations such as (L*) need not be meaningless. In the case at hand, (L*) together with three contemporary certified argumentation repertoires serve as a heuristic to structure the arguments provided by Darwin and his compatriots in defence of 'natural selection'. The three repertoires are labeled 'inductivist', 'demarcationist', and 'professionalist'. Each of these is generated by philosophical or sociological theories of science. Each therefore accounts not only for Darwin's use of certain arguments but also for further historiographical phenomena which are not explained by the theories generating either of the other two repertoires.

ii.

The terms 'induction', 'inductive', and 'inductivist' have been used in a variety of ways. Therefore, any attempt at characterizing the 'inductivist' repertoire of arguments will have to set out with a more general specification of what is meant here by 'inductivist'.

Hume's famous problem of induction can be posed in form of the question: is there truthpreserving ampliative inference?[9] So-called deductivists sometimes impute to so-called inductivists that they are foolhardy enough to answer this question affirmatively. However, the term 'inductivist' is more usefully employed as a label for those philosophers who are concerned with the explication of 'induction' as that term was already used in the second chapter of this study, namely as the third and last stage in Peirce's theory of inquiry. According to that theory, 'abduction' designates the invention of an explanatory hypothesis, 'deduction' its explication and the derivation of testable consequences from it, and 'induction' its evaluation in light of evidence, leading to the rejection of the hypothesis or (quite literally) to its induction into the body of knowledge[10].

> Deduction explicates; Induction evaluates: that is all. Over the chasm that yawns between the ultimate goal of science and such ideas of Man's environment as, coming over him during his primeval wanderings in the forest, while yet his very notion of error was of the vaguest, he managed to communicate to some fellow, we are building a cantilever bridge of induction, held together by scientific struts and ties. Yet every plank of its advance is first laid by Retroduction[11] alone, that is to say, by the spontaneous conjecture of instinctive reason; and neither Deduction nor Induction contributes a single new concept to the structure.[12]

The cantilever bridge of knowledge is evaluated not for the firmness of its 'foundation', but for the quality and expediency of the scientific struts and ties that hold it together. Accordingly, the goal of induction as the final, purely evaluative stage of inquiry is not the establishment of '(absolute) truth' or 'certainty', but the formation and fixation of belief.

> [...] the sole object of inquiry is the settlement of opinion. We may fancy that this is not enough for us, and that we seek not merely an opinion, but a true opinion. But put this fancy to the test, and it proves groundless; for as soon as a firm belief is reached we are entirely satisfied, whether the belief be false or true. [...] The most that can be maintained is that we seek for a belief that we shall *think* to be true. But we think each one of our beliefs to be true, and, indeed, it is a mere tautology to say so.[13]

Now, the most elegant and sweeping 'solutions to the problem of induction' fail in precisely this aspect: they cannot account for the actual formation of belief. Both, Karl Popper and Hans Reichenbach present ways of diminishing the force of Hume's problem. And in order to account for 'belief', both have to invoke principles which they do not (and cannot) justify.

Popper maintains that scientific method can be construed without reference to truthpreserving ampliative inference (i.e. inductive inference). Instead, falseness-transmitting, deductive *modus tollens* is the form of inference characteristic for science. While inductive inference from evidence would presumably justify a hypothesis and *a fortiori* our belief in its truth, *modus tollens* does not contribute to the justification of hypotheses at all[14]. It merely certifies the overall reliability of scientific method: negative instances will falsify hypotheses, the principle of falsification is the cornerstone of scientific method, scientists therefore have to actively anticipate the possibility of such falsification – and to anticipate falsification is to withhold belief[15]. However, in a chapter which is explicitly devoted to "my solution of the problem of induction"[16] Popper acknowledges that to solve Hume's problem, one has to solve it in both its logical and its psychological form, that is, concerning reliable modes of scientific inference as well as the formation of belief.

> H_L Are we justified in reasoning from (repeated) instances of which we have experience to other instances (conclusions) of which we have no experience? [...]
>
> H_{Ps} Why, nevertheless, do all reasonable people expect, and *believe* that instances of which they have no experience will conform to those of which they have experience?[17]

Popper presents his 'solution' to that double-edged problem in a fashion that cleverly oscillates between tongue-in-cheek logical deceptiveness and a serious argument which not so much solves the problem, but forcefully establishes that the logical problem is irrelevant to an analysis of specifically *scientific* reasoning and that the psychological problem can therefore go unanswered as far as the specifics of scientific reasoning are concerned. Two sentences from his autobiography convey this ambiguous use of 'to solve':

> As for induction (or inductive logic, or inductive behaviour, or learning by induction or by repetition or by "instruction") I assert that there is no such thing. If I am right then this solves, of course, the problem of induction.[18]

Indeed, if there is no such thing as induction, the methodology which ensures the progress and rationality of science has to be based on some non-inductive form of inference. By the same token, the issues of justification and belief have to be viewed as irrelevant in respect to that distinctive methodology. But surely, this does not 'solve' the psychological problem of induction that was posed above as 'H_{Ps}'. This scepticism concerning Popper's use of the words 'to solve' and 'solution' persists as one takes a closer look at his somewhat more extensive "Restatement and Solution" of "The Logical Problem of Induction". The overall strategy is clear enough. After providing a solution to H_L, the logical problem, a so-called principle of transference is invoked that will lead to the solution of H_{Ps}.

> *principle of transference*: what is true in logic is true in psychology. (An analogous principle holds by and large for what is usually called 'scientific method' and also for the history of science: what is true in logic is true in scientific method and in the history of science.)[19]

This principle in conjunction with the solution to H_L entails the solution of H_{Ps}. Now, before giving his solution to H_L, Popper reformulates it twice:

> L_1 Can the claim that an explanatory universal theory is true be justified by 'empirical reasons' [...]?
>
> [...] L_2 Can the claim that an explanatory universal theory is true or that it is false be justified by 'empirical reasons' [...]?[20]

L_1 is clearly a reformulation of H_L. And Popper's

> answer to the problem is the same as Hume's: No, it cannot; no number of true test statements would justify the claim that an explanatory universal theory is true.[21]

One should expect that this answer to L_1 is tantamount to the admission that a solution to the problem of induction cannot be provided. Instead, Popper goes on to L_2 which he labels a 'generalization' of L_1 and which accordingly is also proposed as a reformulation of Hume's logical problem of induction. This is where the flippant use of 'to solve' comes in. Popper's positive answer to L_2 cannot be considered a solution or a partial solution to Hume's problem. That positive answer states that

> *the assumption of the truth of test statements sometimes allows us to justify the claim that an explanatory universal theory is false.*[22]

This answer to L_2 cannot be taken as an answer to the problem of induction, since L_2 is not properly a generalization of L_1 and thus a replacement rather than a reformulation of Hume's logical problem H_L. L_2 is supposed to be a generalization, since

> It is obtained from L_1 merely by replacing

> the words 'is true' by the words 'is true or that it is false'.[23]

Now, a proposition Q entails the 'generalization' $(Q \vee R)$ only if Q is true. As there is no affirmative answer to L_1, the affirmative answer to L_2 does not reflect on L_1 and H_L. In effect, Popper has substituted the unsolvable problem of induction by the solvable problem of refutation[24]. Aided by the principle of transference, Popper can show why one would want to disbelieve the truth of a falsified theory. He cannot show why one would ever want to believe that a theory is true[25]. Accordingly, the latter part of the article on the 'solution of the problem of induction' does not rely on the assumption that this solution has been provided by the answer to L_2. Instead of further demonstrating how he has 'solved' Hume's logical problem, Popper proposes an alternative approach towards accounting for belief, that is. On that approach, Hume's psychological problem of induction is taken on directly and without recourse to the principle of transference:

> a *pragmatic belief in the results of science* is not irrational, because there is nothing more 'rational' than the method of critical discussion, which is the method of science. And although it would be irrational to accept any of its results as certain, there is nothing 'better' when it comes to practical action: there is no alternative method which might be said to be more rational.[26]

As stated so far, this alternative approach is clearly insufficient as an answer to H_{PS}. Justifying a commitment to the best available method leaves entirely unresolved when and why any particular result of that method can or should be accepted and a belief should be formed. A possibly appropriate amendment to Popper's alternative account has been proposed (in seemingly quite different a context) by Hans Reichenbach who presented his own 'solution' to the problem of induction. But Reichenbach, too, falls short of justifying the formation of belief and solving the problem of induction.

Though Popper and Reichenbach quite disagree on what is the best method of science, both try to justify actual belief in terms of rational adherence to some such best method. The quality of that method as employed in scientific investigation lends credence to the results of scientific investigation. But while Popper's candidate for 'best method' is founded upon his answer to L_2, Reichenbach's inductivism does not presuppose that one should be able to answer L_1 or any substitute thereof[27]. He argues instead that – if there are any regularities in nature – there are inductive methods which are sufficiently well-suited to the task of finding them. His argument is quite straightforward: One is free to choose whether one wants to consider the world predictable or not. To embark upon the enterprise of science is to reveal at least a tentative preference in this respect, namely to consider it predictable. Now, the term 'predictable' is introduced

> for a world which is sufficiently ordered to enable us to construct series with a limit.[28]

And, it is the aim of induction

> *to find series of events whose frequency of occurence converges toward a limit.*[29]

Now, Reichenbach has to do no more but to show how inductive methods are suited to describe series of events as converging towards a limit of frequency. And his (self-corrective) inductive principle does just that[30]. Then, if the world is predictable, i.e.

> if there is a limit of the frequency, the inductive principle is a sufficient condition to find it.[31]

And conversely, the world is predictable only if inductive methods could find a limit of frequency, i.e.

> the *applicability* of the inductive principle is a necessary condition of the existence of a limit of the frequency.[32]

And therefore, our very commitment to the enterprise of science entails a commitment to the inductive principle. Like Popper's "method of critical discussion", the inductive principle is thought to be the best, if not only avenue to scientific success. Reichenbach realizes, though, that this result does not sufficiently account for the formation of belief in any particular instance. The inductive principle generates a procedure such that the last value for the relative frequency in a series of events is always a candidate for, but never known to be the place of convergence. This procedure has to be employed unintermittently as one seeks the true value of that place of convergence. At any given time, however, one will treat the last available value as the best and most reliable value.

> This procedure must at sometime lead to the true value p, if there is a limit at all; the applicability of this procedure, as a whole, is a necessary condition of the existence of a limit of p. [...] If, however, it is only the whole procedure which constitutes the necessary condition, how may we apply this idea to the individual case before us?[33]

Reichenbach applies it by treating the last value in analogy to a "blind posit" in a wager.

> We know it is our best posit, but we do not know how good it is. Perhaps, although our best, it is a rather bad one. The blind posit, however, may be corrected. [...] Thus the blind posit is of an approximative type; we know that the method for making such posits must in time lead to success, in case there is a limit of the frequency. It is this idea which furnishes the justification of the blind posit. The procedure described may be called the *method of anticipation*; in choosing h^n as our posit, we anticipate the case where n is the "place of convergence." It may be that by this anticipation we obtain a false value; we know, however, that a continued anticipation must lead to the true value, if there is a limit at all.[34]

There is a fundamental ambiguity in this passage: the idea that the method of making blind posits must in time lead to success may justify the making of blind posits, it does not justify, however, the posits themselves. From the point of view of the inductive principle the blind posit that one actually makes or uses has no special standing. All one has to do is to keep the procedure going and to always be prepared to throw out what was considered the best value and substitute it by a newly found last value. That is, one can make blind posits as long as one does not believe and does not think that the last value in a series of events is the true value or place of convergence[35]. Thus, Reichenbach does not answer his question: the method of anticipation does not account for the formation of belief in a given instance and the principle of induction is not applied to "the individual case before us".

Instead of demonstrative inductive inference, one has according to Popper and Reichenbach demonstrably sound methods for the production of knowledge - those, however, do not by themselves justify belief in any given instance. Proposing such demonstrably sound methods at the expense of accounting for the formation of belief is a way of circumventing rather than addressing Hume's (or Peirce's) problem of induction, quite regardless of the merits of the specific proposals. One may surmise that Popper's and Reichenbach's fault lies in their attempt to provide a wholesale demonstration which would justify belief in all results obtained by adherence to a given method. To be tried instead is a rather more myopic approach which provides for a case-by-case evaluation in respect to induction. And in contradistinction to Popper and Reichenbach, the goal of the inductivist shall be described as specifying the conditions under which one is warranted to accept a hypothesis as true *in the absence of demonstrable justification of its truth.* This is in agreement with Peirce's notion of 'induction' as the evaluation of a hypothesis for the purpose of induction into the body of knowledge whether or not one can justifiably hold it to be true. More succinctly put, inductivists are not concerned with the justification of hypotheses, whether directly (by inductive inference) or indirectly (owing to the soundness of method). Instead, they are concerned with human decision-making: at any given time one decides on the best ranking of preferences among the options of believing a hypothesis, disbelieving it, or suspending judgement on it - where 'to believe', 'to disbelieve', and 'to suspend judgement' form an ultimate partition, i.e. an exclusive and exhaustive set of options[36]. In relation to this larger enterprise of inductivism, all attempts at specifying measures for e.g. degrees of confirmation are subservient contributions of detail, and as of now neglectable at that. If successful, specification of such measures may provide *one* important kind of input into the evaluative process[37].

Induction conceived of as the formation of a preference within an ultimate partition of theoretical attitudes towards hypotheses, is a species of what was formerly called 'eliminative' in contradistinction to 'enumerative induction'. A complete account of what goes into this process of evaluation, preference-ranking and elimination of alternatives cannot be given here. A somewhat sketchy presentation of the necessary tools of devices driving the inductive machinery must suffice[38]. *Background knowledge* provides the *standard of serious possibility*. It outrules hypotheses which contradict presently held beliefs: one can seriously entertain only hypotheses which are compatible with one's present background knowledge. A *demand for information* generates an ultimate partition of potential answers to a given problem. (This ultimate partition is not to be confused with the just-mentioned ultimate partition of potential theoretical attitudes towards any of these answers.) The informational value of the potential answers is assessed as one applies an *information-ranking function*. For instance, one might consider "theories allowing for action at a distance to bear higher [...] informational value than those which insist on contact action" or *vice versa*[39]. Once a preliminary ranking of potential answers has been established, evidence is brought to bear on them. With the introduction of evidence, the answers require re-evaluation. Here, *confirmational commitments* determine the relative weight one decides to assign to various kinds of confirmatory or falsificatory evidence. Rankings according to informational value and evidential support subsequently enter into the larger task of weighing the *risk of error* against the *benefit of informational gain* in respect to the ultimate partition of theoretical attitudes, i.e. believing, disbelieving or suspending judgement.

The conceptual apparatus presented so far closely conforms to Peirce's insight that 'belief' is tantamount to 'settled opinion' on the one hand, to 'acceptance as true' on the other. However, the role assigned to background knowledge as the standard for serious possibility points to what is apparently a serious defect in Peirce's formulation. While equating 'settled opinion' with 'acceptance as true' agrees entirely with experience and intuition as far as it captures the feeling of confidence associated with 'belief', it is also counter-intuitive insofar as it does not seem to provide for the possibility of changing one's mind or of having a sceptical or critical attitude in general. But there is indeed a rather simple way of remedying that defect. All that is required is a proper mode of sequencing the decision-making situations which constitute the phenomenon 'change of mind'[40].

The induction of a newly formed belief into a corpus of knowledge *expands* that corpus. One can distinguish 'routine expansion' and 'inferential expansion'[41]. While inferential expansion is the result of more or less explicit deliberation, routine expansion generally ensues upon some antecedent commitment to admit certain kinds of new information into the corpus, for instance the information gained through sensory stimulation or the testimony of trusted friends. Once the corpus is expanded, the freshly inducted belief constitutes a standard of serious possibility. It may happen now that conflict arises as one imports a contradiction into one's corpus by routine or inferential expansion. In response to this conflict, the corpus will be *contracted* and what was formerly a firmly held belief is contemplated once again for its cogency. Contraction, just like expansion, takes place in response to evidence: one needs grounds for doubting as much as for believing. After contraction one is free to entertain alternatives which previously were not considered seriously possible. One may be in suspense between various accounts for any length of time, until by inferential expansion a new or modified version of the eliminated belief is inducted into the corpus.

Without compromising the notion that belief is a standard for serious possibility at any given time (and thus considered infallible), revision of belief can be represented by the sequence of expansion, contraction, and re-expansion of a corpus of knowledge. But aside from integrating infallibilism and corrigibilism, the sequential reconstruction of 'change of belief' suggests a mode of accounting for formation and change of belief in science. Indeed, an earlier episode from Charles Darwin's scientific career allows for a paradigmatic application of these conceptual tools[42]. In 1839 Darwin published a paper on the "Parallel Roads of Glen Roy"[43]. In it, he argued for the adoption of a hypothesis on the origin of a puzzling geological phenomenon. He had convinced himself of its truth in a series of steps. First, he had drawn up an ultimate partition of potential answers to the problem. That ultimate partition contained a number of previously proposed hypotheses and Darwin's own suggestion which was part of a larger geological research programme. He then applied an information-ranking function, namely Lyell's actualism: on the regulative principle of actualism, highest informational value is accorded to those hypotheses which make reference only to causes which are now (and always) in operation[44]. From this application emerged his own proposal as the preferred hypothesis. Only then he traveled to the site and collected evidence in support of that hypothesis. Though the empirical evidence was not as clearly in favor of Darwin's theory as he expected it to be, he satisfied his confirmational commitment by detecting a number of phenomena which corroborated his view and which were inexplicable on the alternative accounts. The published paper presented his empirical findings as part of an argument which eliminated all the rivaling hypotheses from the ulitmate partition, leaving - on the 'principle of exclusion' - his own as the only viable account. Some time after the induction of his geological hypothesis into his corpus of belief he came by routine expansion upon another hypothesis which had not been part of the ultimate partition. Moreover, the new proposal met the informational demands imposed by Lyell's actualism. Believing as he did, that he had correctly applied the pinciple of exclusion on an exclusive and exhaustive set of potential answers, getting to know an overlooked hypothesis was tantamount to importing a contradiction into his corpus of belief. Darwin immediately realized the full impact of his oversight. He contracted his corpus and remained in anxious suspense among the two preferred hypotheses. After suspending judgement for just about 21 years, Darwin finally rejected his proposal and accepted the remaining rival.

As the inductivist would predict and as Martin Rudwick and David Hull point out, Darwin behaved as a rational agent throughout this process. The story of Darwin's hypothesis on the Parallel Roads of Glen Roy tells of no failure in scientific method, on the contrary, it tells of ultimate success. Darwin prudently weighed the risk of error which is inevitably linked to drawing up an ultimate partition of potential answers, against informational gain and evidential support. In effect, he did so twice. For, as he finally chose to accept the remaining rival to his own theory, he again selected the best among an ultimate partition of potential answers. If Darwin's theory on Glen Roy has nevertheless been considered "a long gigantic blunder from beginning to end"[44], it is because Darwin himself called it so.

> Having been deeply impressed with what I had seen of the elevation of the land of South America, I attributed the parallel lines to the action of the sea; but I had to give up this view when Agassiz propounded his glacier-lake theory. Because no other explanation was possible under our then state of knowledge, I argued in favour of sea-action; and my error has been a good lesson to me never to trust in science to the principle of exclusion.[46]

Here, Darwin provides a strangely ambiguous assessment. On the one hand he clearly justifies his procedure: given the constraint of the "then state of knowledge" he finds that, indeed, "no other explanation was possible". On the other hand, he says that he has learned to distrust the principle of exclusion. Both, Rudwick and Hull produce a coherent interpretation of these statements by specifying precisely what Darwin may have learned to distrust.

> What was at fault was not the "principle of exclusion" itself, but the *degree* of Darwin's trust in it. In other words, had he been more

> cautious, he would have inserted a proviso: "Since A cannot account for the observed phenomena, while B can, B must be the "true cause" -- *unless* there are other alternatives, C... etc., which I have not considered."[47]
>
> Darwin was correct to doubt this principle but wrong in thinking that he could do without it. He continued to use the principle of exclusion throughout his later works. What he had learned was not that the principle of exclusion could *never* be trusted but that it could never be trusted *completely*. [...] It could not in actual practice make a hypothesis absolutely certain, since in the natural sciences it is never possible to eliminate *all* alternatives.[48]

The latter statement suggests not only that Darwin continued using that principle, but that furthermore - as the inductivist would predict - science cannot do without it. After all, the inductivist problem of accounting for the formation of belief became pressing precisely as the impossibility of ampliative inductive inference led to a shift from the evasive ideal of 'certainty' to 'well-reasoned preference'.

The inductivist repertoire controls the application of the principle of exclusion. It concerns the identification of a demand for information, the construction of an ultimate partition of potential answers, the choice of an information-ranking function over the members of the partition, and finally the confirmational commitments which explicate the evidential grounds for preferring some one among all potential answers[49]. While scientists use that repertoire to justify their belief that a certain choice among alternatives in science is the right choice to make, scientists of science can use it to explain that:

> (Ph3) Scientists form beliefs, i.e. tend to identify preferred theories with true theories.[50]

Since a belief is a standard for serious possibility, an explanation of (Ph3) can contribute to an understanding not only of consensus and paradigm formation but more specifically of tenacity in science: once a belief is formed, scientists need not take seriously newly proposed rivals until a point is reached where by further ignoring the rival one would import a contradiction into one's corpus - either, as in Darwin's case, a contradiction with current beliefs (e.g. on the correctness of an ultimate partition) or a violation of an internal standard of reasoning[51].

From the vast resource of material provided by Darwin and his contemporaries, only very few statements shall now be used to briefly illustrate the workings of the inductivist repertoire towards the formation of Darwin's belief in 'natural selection'.

> In considering the Origin of Species, it is quite conceivable that a naturalist, reflecting on the mutual affinities of organic beings, on their embryological relations, their geographical distribution, geological succession, and other such facts, might come to the conclusion that each species had not been independently created, but had descended, like varieties, from other species. Nevertheless, such a conclusion, even if well founded, would be unsatisfactory, until it could be shown how the innumerable species inhabiting

> this world have been modified, so as to acquire that perfection of structure and coadaptation which most justly excites our admiration. [...] The author of the 'Vestiges of Creation' would, I presume, say that, after a certain unknown number of generations, some bird had given birth to a woodpecker, and some plant to the misseletoe, and that these and been produced perfect as we now see them; but this assumption seems to me to be no explanation, for it leaves the case of the coadaptations of organic beings to each other and to their physical conditions of life, untouched and unexplained. It is therefore of the highest importance to gain a clear insight into the means of modification of coadaptation.[52]

This passage identifies a demand for information, namely, to find an explanatory scheme for the process of evolution. And already, it applies a most elementary information-ranking function to clear the field of such potential answers as the one presumably provided by Chambers. Indeed, Darwin denies that Chamber's theory was abductively successful which surely rules it out as a potential answer.

Darwin's answer remained without serious rival. The ultimate partition of potential answers consisted only of Darwin's proposal and a residual hypothesis which merely states that there may be other solutions[53]. The poverty of that ultimate partition allowed Darwin to apply a simple and uncontroversial information-ranking function which really stated no more than: if there is demand for information and if it is met by only one theory, then its informational valued and usefulness should be ranked considerably higher than that of the vacuous residual hypothesis. Only therefore, the explanatory value or abductive success of the theory could be used as an argument for its acceptance:

> I have always looked at the doctrine of Natural Selection as an hypothesis, which, if it explained several large classes of facts, would deserve to be ranked as a theory deserving acceptance; and this, of course, is my own opinion.[54]

To be sure, on occasion Darwin also employed the information-ranking which he had used at Glen Roy.

> I further believe, that this very slow, intermittent action of natural selection accords perfectly well with what geology tells us of the rate and manner at which the inhabitants of this world have changed.[55]

Though reference to Lyell's actualism may have been designed to persuade Lyell himself and other Lyellians, Darwin did not actually need to invoke this principle to argue for the (tentative) acceptance and belief of a theory which is an only serious contender:

> I have now at last satisfied *myself* (but that is very different from satisfying others) on this head [...] I dare say you will think all this utter bosh, but I believe it to be solid truth.[56]

- iii.-

Darwin drew upon the inductivist argumentation repertoire in order to justify his belief in the truth of the theory of evolution by natural selection. In his review on the *Origin* Richard Owen criticized what he took to be the exclusive employment of only this argumentation strategy:

> Now, on such a question as the origin of species, and in an express, formal, scientific treatise on the subject, the expression of a belief, where one looks for a demonstration, is simply provoking. We are not concerned in the author's beliefs or inclinations to believe. Belief is a state of mind short of actual knowledge. It is a state which may govern action, when based upon a tacit admission of the mind's incompetency to prove a proposition, coupled with submissive acceptance of an authoritative dogma, or worship of a favourite idol of the mind.[57]

While Owen does not deny that Darwin may have good reasons for believing his theory to be true, he insists that any expression of belief in the absence of demonstration or proof plays no significant role in a scientific argument. After all, Owen himself may have good reasons for disbelieving that theory which would lead to unproven conviction pitched against unproven conviction. And Darwin, as he reports to Lyell a conversation with Owen, appears ready to agree with him:

> He added: - "If I must criticise, I should say, we do not want to know what Darwin believes and is convinced of, but what he can prove." I agreed most fully and truly that I have probably greatly sinned in this line, and I defended my general line of argument of inventing a theory and seeing how many classes of facts the theory would explain. I added that I would endeavour to modify the "believes" and "convinceds." He took me up short: "You will then spoil your book, the charm of (!) it is that it is Darwin himself."[58]

Darwin's readiness to concede that the expression of his convictions and beliefs should be modified, corresponds to a shift from one argumentation repertoire to another, namely from the inductivist repertoire to one that is here construed as 'demarcationist'[59].

In the preceding quotations, a proof or demonstration was thought to provide a compellingly forceful scientific argument. And regardless of whether there actually are proofs and demonstrations in the natural sciences and particularly in biology, a scientist's conviction or belief is thought to be entirely circumstantial or epiphenomenal as far as specifically scientific reasoning is concerned[60]. Owen and Darwin are thus expressing Popper's view on the role of belief in science.

As was shown above, Popper's 'solution' to the problem of induction does not include an account of how belief is formed and when one may be justified in believing a theory to be true[61]. Rather, it provides a way of circumventing or bypassing that problem. Belief is epiphenomenal to the scientific enterprise insofar as one may rationally believe all sorts of things, e.g. hypotheses which contain mystery-raising predicates as well as unproblematic theories[62]. Unless one can positively prove or demonstrate the truth or falsity of a scientific theory, showing one's grounds for believing it is no substitute for establishing that it can be subjected to meaningful scientific criticism. Informed by Popper's *critical* rationalism as the hallmark of science, the demarcationist argumentation repertoire caters to the methodological posture which scientists adopt as they argue that a given theory is properly scientific, that it raises problems rather than mysteries.

Popper's methodology is based on the now largely uncontroversial recognition that there is no positive proof or demonstration in science, but at best negative proof or falsification. It is methodologically irrelevant that scientists do indeed positively believe in the truth of theories. It is equally irrelevant that they do so on rational grounds, i.e. in accordance with inductive rules on decision under unresolved conflict. Their belief is, so to speak, a by-product of scientific work, a dimension of human decision-making rather than of evaluation on characteristically scientific principles. Even if one maintains that this by-product becomes an indispensable force in shaping the further development of a theory and of science, the fundamental difference of orientation between Popper's demarcationist and Carnap's inductivist programme subsists. That difference is clearly expressed by Stegmüller:

> There are absolutely no points of contact between Popper's and Carnap's theories. [...]

> Popper's theory of corroboration pertains to the *theoretical evaluation of unverifiable hypotheses*. Carnap's theory pertains to the *establishment of norms for human decisions under risk*[63].

A human decison under risk or unresolved conflict does not settle theoretical problems: in regard to a scientific theory, conflict *remains* unresolved and everything is open unless that theory has been falsified. To be sure, theorists of science and scientists themselves need to integrate the inductivist posture which gears argumentation towards persuasion and the demarcationist posture according to which belief is altogether irrelevant in science[64]. This integration has to take place on a different level of scientific reasoning which will be introduced in the following section with the 'professionalist' repertoire. For now, from the Popperian point of view it clearly suffices to present the following sort of argument: If the professed open-mindedness of science appears disingenuous as far as the private person, the individual scientist is concerned, it is precisely this deliberately cultivated and sustained disingenuousness which distinguishes the collective engagement in the search for scientific truth. After all, a public commitment to steadfast open-mindedness sustains certain philosophical, cultural, and political norms and values quite regardless of whether there is a deeply felt belief in these norms and values on a personal level. What is true on the level of cultural norms holds also for the norms of conduct prescribed by the methodologically sound scientific procedure of falsificationism: aside from everything else scientists may think and do, they have to ensure that they *also* conform to the falsificationist standard of critical inquiry which legitimizes their work methodologically (and possibly also socially or culturally[65]).

It is therefore not enough for Darwin to address a demand for information, to propose a theory which meets that demand, and then to argue that it meets the demand in such a way that for any number reasons it should be accepted. Darwin had to show moreover that his theory meets the demand for information in a particular way, namely in a way that makes it available to critical discussion and thus worthy of scientific concern[66]. For this purpose, a mystery-raising attitude towards it had to be outruled and a firm establishment of its empirical significance had to be pursued. The demarcationist argumentation repertoire is geared towards that end: falsifiability is a criterion that demarcates science from pseudo-science and thus the problem-raising from the mystery-raising attitude. As such, the repertoire first and foremost includes all those arguments which address the questions whether a theory is falsifiable or not and, if yes, whether or not it has actually been falsified so far.

However, this is not the full extent of the Popperian repertoire. Firstly, it encompasses arguments aimed at the theoretical evaluation of competing scientific theories in terms of their past performance in respect to the falsifiability requirement (one might call this the regulatory extension of the demarcationist repertoire). And more importantly for the case at hand, it also includes a heuristic extension for the theoretical evaluation of non-scientific, metaphysical, or not yet certifiably scientific ideas in respect to the falsifiability requirement. Both these extensions involve the notion of 'theoretical evaluation'. Apparently this notion has been much misunderstood and should be clarified first - especially since Popper himself fuels that misunderstanding when he presents in the following terms the regulative extension of the repertoire:

> Of course theories which we claim to be no more than conjectures or hypotheses need no justification (and least of all a justification by a nonexistent "method of induction", of which nobody has ever given a sensible description). We can, however, sometimes give reasons for preferring one of the competing conjectures to the others, in the light of their critical discussion.[67]

More specifically, a theory's ability to withstand attempts at falsification is to provide grounds for such preference.

> Although we cannot justify a theory - that is, justify our belief in its truth - we can sometimes justify our *preference* for one theory over another; for example if its degree of

corroboration is greater.[68]

The notion that a theory's corroboration can serve as a reason for preferring it over other theories has been taken as a none too well disguised re-issue of inductive confirmation, and Popper was seen to be inadvertently slipping from falsificationism to inductivism[69]. Stegmüller vigorously defends Popper on this score:

> Popper's proposal states: in each domain of empirical inquiry one is to choose the *empirically most significant* theory among all the rivaling theories - that is the one *with the most potential falsificatory instances*, i.e. the most risky and *logically least probable one*. - And all of this not for the purpose of accepting it but in order *to subject it to strict testing* (more precisely: in order to subject it to the strictest test which one can think of).[70]

Stegmüller's clarification suggests the source of the misunderstanding. Where Popper juxtaposes 'belief' and 'preference', Stegmüller simply speaks of 'acceptance'. And while Popper seems to be saying that one cannot scientifically justify belief in the truth of a theory but that one can sometimes justify one's preference for one theory over all other known theories, Stegmüller bypasses this presumed distinction by simply denying that one can scientifically justify the acceptance of a theory. Indeed, Stegmüller's reading is the only plausible construal of Popper's argument. Surely Popper would not reject 'belief' as a purely epiphenomenal internal state of mind of individual scientists only to posit 'preference' as another, supposedly distinct and somehow scientifically justifiable state of mind[71]. And indeed, in the *Logic of Scientific Discovery* Popper made the same point as above without invoking the notion of 'preference'.

> Scientific theories can never be 'justified', or verified. But in spite of this, a hypothesis *A* can under certain circumstances achieve more than a hypothesis *B* – perhaps *B* is contradicted by certain results of observations, and therefore 'falsified' by them, whereas *A* is not falsified; or perhaps because a greater number of predictions can be derived with the help of *A* than with the help of *B*. The best we can say of a hypothesis is that up to now it has been able to show its worth, and that it has been more successful than other hypotheses although, in principle, it can never be justified, verified, or even shown to be probable. This appraisal of the hypothesis relies solely upon *deductive* consequences (predictions) which may be drawn from the hypothesis. *There is no need even to mention induction.*[72]

'Corroboration' or 'degree of corroboration' expresses the resilience of hypotheses in the face of attempts at falsification. Faced with a well-corroborated hypothesis, i.e. in the absence of grounds for its rejection, scientists stick with it *nolens volens*. They 'prefer' a well-corroborated and highly falsifiable theory only insofar as they follow a rule which demands that resilient scientific theories with considerable empirical content are first on the list of theories requiring attempts at falsification. The relevant distinction is not between 'belief' and 'preference' as assumed states of mind, but between 'preference by rule' and 'preference as human choice'. 'Preference by rule' is a species of theoretical evaluation resulting from the application of a rule that operates on the degree of confirmation and which has somehow been deduced from falsificationist demands on scientific theories. To be sure, Popper himself has not formulated such a rule:

> I do not propose any 'criterion' for the choice of scientific hypotheses: every choice remains a risky guess. Moreover, the theoretician's choice is the hypothesis most worthy of *further critical discussion* (rather than *acceptance*).[73]

All negotiation on the theoretician's elusive "most worthy", i.e. all attempted specifications of that preference-guiding rule belong to what has above been termed the 'regulatory extension' of the demarcationist repertoire[74]. Thus, 'preference by rule' in contrast to 'preference as human choice' is construed in strictly deductive manner, while all human choice including the formation of any state of mind (preferring or believing as well as disbelieving) involves an act of induction.

The theoretical evaluation of rivaling scientific theories surely requires elaborate negotiations on notions like 'degree of corroboration', 'comparative predictive power', or 'relative progressiveness of research progammes'. Moreover, questions of what is epistemologically sound procedure in respect to any of these notions will arise, and these, in turn, are frequently discussed in more general philosophical terms, e.g. as critically employed metaphysics[75]. But as difficult as this process of negotiation may be, it stills seems far more tangible than the process of negotiation on the heuristic value of non-scientific or not-yet certifiably scientific theories. For, in the latter case, one has to anticipate the scientific value of metaphysical ideas[76].

While both Popper and Imre Lakatos acknowledge the science-generative force of many metaphysical ideas, it took Lakatos's notion of a 'positive' and 'negative heuristic' to suggest if only weak criteria for the theoretical evaluation of metaphysical core-ideas:

> All scientific research programmes may be characterized by their *'hard core'*. The negative heuristic of the programme forbids us to direct the *modus tollens* at this 'hard core'. Instead, we must use our ingenuity to articulate or even invent 'auxiliary hypotheses', which form a *protective belt* around this core, and we must redirect the *modus tollens* to *these*. It is this protective belt of auxiliary hypotheses which has to bear the brunt of tests and get adjusted and re-adjusted, or even completely replaced, to defend the thus-hardened core.[77]

While the negative heuristic simply consists of the directive to deflect criticism of the hard core to its protective belt, the construction and modification of the protective belt is part of the positive heuristic.

> 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, the 'refutable' protective belt. [...] One may formulate the 'positive heuristic' of a research programme as a 'metaphysical' principle. [...] We may appraise research programmes, even after their 'elimination', for their *heuristic power*: how many new facts did they produce, how great was their 'capacity to explain their refutataions in the course of their growth'?[78]

Almost paradoxically, then, one would have to say that while the hard core generates scientific work, it is not itself a primary subject of scientific concern. As far as the requirements of the demarcationist repertoire are concerned, critical reasoning and science itself take place in the negotiation on the protective belt which impacts rather obliquely on the hard core. And yet, the overall research programme can be assessed in respect to the requirement of falsifiability which is to be met at least by the auxiliary theories of the protective belt. The demarcationist repertoire contains the arguments used to establish falsifiability for the protective belt as it is given now or as it may develop under the guidance of the positive heuristic. Argumentation from the demarcationist repertoire may thus start off by first acknowledging that a given theory is not or not yet falsifiable in a significant sense (though it should be)[79]. It then proceeds to make the case that nevertheless it would be premature to discard that theory as mystery-raising. In support of this, this line of argumentation may continue by emphasizing that judgement-calls are needed here. And as input to these judgement-calls, falsificationist considerations on the general architecture of theories in terms of beauty and simplicity will be introduced. And possibly in reference to past experience in the history of science, one will argue that the requirement of falsifiability or independent testability is going to be met at some future time.

Especially these latter sorts of arguments from the heuristic extension of the demarcationist repertoire were employed in the negotiation on Darwin's theory of evolution by natural selection[80]. For, the testability of Darwin's theory was the central issue from the falsificationist point of view. Once independent testability has been established, recourse to the heuristic extension is not warranted anymore. And if independent testability cannot be established in principle, recourse to the demarcationist repertoire will do no good. But a quick survey of the issue shows that these matters were quite open in the case of 'natural selection'.

By providing an explanatory scheme for the evolutionary process, the theory of natural selection created a protective belt around the theory of evolution which constitutes the hard core of the research programme 'Evolutionary Biology'. As such, the theory of natural selection insulated the theory of evolution from immediate criticism. Indeed, the negotiation of natural selection *requires* at least a tentative acceptance of evolution: insofar as the theory of natural selection provides an explanatory mechanism for evolutionary processes, the merits of that explanatory theory can be assessed only once it is agreed upon that there are evolutionary processes which ought to be accounted for. Simply by introducing the theory of natural selection, Darwin shifted the status of evolution radically from daring biological conjecture to theoretically embedded state of affairs. This was most clearly expressed by George Romanes in 1896:

> we must have some reasonable assurance that a fact is a fact before we endeavour to explain it. Nevertheless, it is not necessary that we should actually demonstrate a fact before we endeavour to explain it. Even if we have but a reasonable presumption as to its probability, we may find it well worth while to consider its explanation; for by so doing we may obtain additional evidence for the fact itself. And this because, if it really is a fact, and if we hit upon the right explanation of it, by proving the explanation probable, we may thereby greatly increase our evidence of the fact. In the very case before us, for example, the evidence of evolution as a fact has from the first been largely derived from testing Darwin's theory concerning its method. It was this theoretical explanation of its method which first set him seriously to enquire into the evidences of evolution as a fact; and ever since he published his results, the evidences which he adduced in favour of natural selection as a method have constituted some of the strongest reasons which scientific men have felt for accepting evolution as a fact. Of course the evidence in favour of this fact has gone on steadily growing, quite independently of the assistance which was thus so largely lent to it by the distinctively Darwinian theory of its method[81]

Romanes's somewhat cumbersome treatment concludes with a euphemism. Where he notes that evolution was adopted 'quite independently' of the theory of natural selection which precipitated its adoption, one should almost say 'in spite of'. The theory of natural selection steadily lost support towards the end of the century while the theory evolution rather swiftly became generally accepted doctrine. The relation between accepted hard core and tentative protect belt can thus be used to explain phenomenon (Ph4).

> (Ph4) Darwin's contribution consisted of providing an explanatory scheme for the controversial theory of evolution. He did not introduce significant, direct evidence for the process of evolution. And yet, as an immediate consequence of his work, the theory of evolution was swiftly accepted, while the theory of natural selection was for a long time considered highly doubtful[82].

In strictly Popperian terms, then, insofar as the core of the research programme of evolutionary biology cannot be criticized directly, it is by definiton 'metaphysical':

> It is metaphysical because it is not testable. [...] For assume that we find life on Mars consisting of exactly three terrestrial species. Is Darwinism refuted? By no means. We shall say that these three species were the only forms among the many mutants which were sufficiently well adapted to survive.[83]

Evolutionary theory can be saved by ad-hoc adjustments, whatever the presenting phenomena. These ad-hoc adjustments, however, are generated from the evolutionary core by the theory of natural selection on the protective belt. Since that latter theory may well be (independently) testable, each ad-hoc adjustment that is designed to save the evolutionary core may have testable consequences and may have to satisfy coherence conditions which severely constrain the scope of ad-hocness. Thus, as Popper continues the passage just cited, he infers somewhat too hastily that the metaphysical core of evolutionary biology renders the entire research programme metaphysical.

> Thus Darwinism does not really *predict* the evolution of variety. It therefore cannot really *explain* it. [...] In other fields, its predictive or explanatory power is still more disappointing. Take "adaptation". [...] Adaptation or fitness is *defined* by modern evolutionists as survival value, and can be measured by actual success in survival: there is hardly any possibility of testing a theory as feeble as this.[84]

Popper's mistake lies in the demand that *each and all* explanatory relations between protective belt and hard core should be accessible to independent test[85]. But whether or not one can ultimately find a way of determining (degrees of) 'adaptation' or 'fitness' independently of actual survival values, the theory of natural selection involves a great number of *other* empirical phenomena, some of which may well constitute independent testability[86]. Indeed, Popper later revised his opinion on the testability of natural selection.

> I still believe that natural selection works in this way as a research programme. Nevertheless, I have changed my mind about the testability and the logical status of the theory of natural selection

> [...] In its most daring and sweeping form, the theory of natural selection would assert that *all* organisms, and especially *all* those highly complex organs whose existence might be interpreted as as evidence of design and, in addition, *all* forms of animal behaviour, have evolved as the result of natural selection; that is, as the result of chance-like inheritable variations, of which the useless ones are weeded out, so that only the useful ones remain. If formulated in this sweeping way, the theory is not only refutable, but actually refuted. For *not all* organs serve a *useful* purpose: as Darwin himself points out, there are organs like the tail of the peacock, and behavioural programmes like the peacock's display of his tail, which cannot be explained by their *utility*, and therefore not by natural selection. Darwin explained them by the preference of the other sex, that is, by sexual selection. Of course one can get round this refutation by some verbal manoeuvre: one can get round any refutation of any theory. But that one gets near to rendering the theory tautological. It seems far preferable to admit that *not* everything that evolves is *useful*, though it is astonishing how many things are; and that in conjecturing what is the *use* of an organ or a behavioural programme, we conjecture a possible explanation by natural selection[87]

Aside from the question whether Popper is right on questions of detail or interpretation[88], this passage is most illuminating in three respects. Firstly, what it says about natural selection is easily restatable in the following now-familiar terms: the theory of natural selection is tautological only if one adopts a mystery-raising attitude towards it. But it can function successfully as a research programme if one continues the negotiation by conjecturally proposing empirical interpretations of that theory, i.e. if one adopts a problem-raising attitude towards it. Secondly, for a long time there was little or no tangible success in the negotiation on natural selection. Indeed, one might infer from Popper's discussion that even today few inroads have been made towards the clarification of natural selection and its empirical significance. But then again, in the absence of a serious contender to the theory of natural selection, the rate of success in the negotiation is not very important as long as the problem-raising attitude is actively pursued. For, thirdly, *in the meantime* natural selection does a perfect job at insulating the theory of evolution.

In accordance with the three remarks on Popper's interpretation of natural selection, one can now identify the following types of arguments which were employed by Darwin in defence of his theory and which were drawn from the demarcationist argumentation repertoire (in its heuristic extension). Into that repertoire belong

- all arguments aimed at establishing natural selection as a protective belt that delineates a proper domain of scientific discourse (leading to the tentative acceptance of evolution),

- all references to the beauty, simplicity, abductive success or explanatory power of the theory, since those references are indices of high empirical content and greater falsifiability,

- all admissions of the theory's failure to presently meet the scientific standard of falsifiability,
- all allusions to analogous episodes in the history of science where a judgement in favor of conceptually difficult theories constituted scientific progress, and
- all anticipations of conditions under which the theory may be or may become falsifiable.

The introduction of the theory of natural selection created a realm of scientific discourse which rendered the theory of evolution acceptable. That is, by *inventing a theory*, Darwin staked out new grounds for scientific research. In the course of negotiation on whether his theory was properly scientific, 'science' was redefined to now include the previously untenable (and untenably ideological) idea of evolution. To quote two contemporary reviews:

> to raise this hypothesis of creations by variation to the rank of a scientific theory, is the object of Mr. DARWIN'S book, and to do this he has endeavoured to invent and prove such an intelligible rationale of the operation of variation, as will account for many species having been developed from a few in strict adaptation to existing conditions[69]

> The history of science shows that the great epochs of its progress are those not so much of new discoveries of *fact*, as those of new *ideas* which have served for the colligation of facts previously known into general principles, and which have thenceforward given a new direction to inquiry. It is in this point of view that we attach the highest value to Mr. Darwin's work.[70]

Surely, it was not enough to simply invent a theory. A theory which will successfully serve as the protective belt for another, has to meet *some* requirements which should enable it to digest the brunt of criticism. In properly Popperian fashion, H. Fawcett, one of Darwin's compatriots presented the following sketch:

> it should at once be distinctly stated that Mr. Darwin does not pretend that his work contains a proved theory, but merely an extremely probable hypothesis. The history of science abundantly illustrates that through such a state of hypothesis all those theories have passed which are now considered most securely to rest on strict inductive principle. [...] The mode therefore is plainly indicated by which the incorrectness of Mr. Darwin's speculations can be completely established. If the physical philosopher can demonstrate that the geological record has not this character of extreme imperfection, Mr. Darwin will doubtless be amongst the first to admit that his theory can then be no longer maintained.[91]

Fawcett's claim that Darwin's theory is falsifiable was rather singular. Indeed, he was speaking here about what is probably a further protective belt: the imperfection of the geological record insulates the theory of natural selection from certain sorts empirical criticism. Moreover, the claim about the imperfection of the geological record is not as easily falsifiable as Fawcett seemed to think.

In contrast, the strategy of Darwin and his followers did not rest on the claim that his theory is falsifiable. And yet, they also argued from a falsificationist (or demarcationist) point of view as they tried to pre-empt falsificationist criticism:

> I am actually weary of telling people that I do not pretend to adduce direct evidence of one species changing into another, but that I believe that this view in the main is correct, because so many phenomena can be thus grouped together and explained.[72]

Accordingly, Darwin was quite delighted about John Stuart Mill's assessment of his theory, and rightly so[73]:

> Mr. Darwin's remarkable speculation on the *Origin of Species* is another unimpeachable example of a legitimate hypothesis. What he terms "natural selection" is not only a *vera causa*, but one to be proved capable of producing effects of the same kind with those which the hypothesis ascribes to it: the question of possibility is entirely one of degree. It is unreasonable to accuse Mr. Darwin (as has been done) of violating the rules of Induction. The rules of Induction are concerned with the conditions of Proof. Mr. Darwin has never pretended that his doctrine was proved. He was not bound by the rules of Induction, but by those of Hypothesis. And these last have seldom been more completely fulfilled. He has opened a path of inquiry full of promise, the results of which none can foresee. And is it not a wonderful feat of scientific knowledge and ingenuity to have rendered so bold a suggestion, which the first impulse of every one was to reject at once, admissible and discussible, even as a conjecture?[74]

This assessment would provide the theory a chance for growth. Marking it as a product of 'hypothesis' or 'abduction' (with 'deduction' and 'induction' still outstanding), has a similar effect as labeling 'natural selection' a metaphor: prudent restraint in the assessment of its scientific standing yields a gain in scientific credibility to the theory and its authors. It shows that they are embarked on a problem-raising course and unwilling to confuse a good explanatory hypothesis with a proven scientific theory. While he kept waiting for 'direct evidence' and with the example of great science in mind, Darwin knew that not all depends on the immediate falsifiability of 'natural selection'.

> It can hardly be supposed that a false theory would explain, in a satsifactory manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers. The undulatory theory of light has thus been arrived at; and the belief in the revolution of the earth on its own axis was until lately supported by hardly any direct evidence. It is no valid objection that science as yet throws no light on the far higher problem of the essence or origin of life. Who can explain what is the essence of the attraction of gravity? No one now objects to following out the results consequent on this unknown element of attraction; notwithstanding that Leibnitz formerly accused Newton of introducing "occult qualities and miracles into philosophy."[95]

To be sure, by far not all of this convoluted argument could have been drawn from the demarcationist repertoire, especially not the first sentence. But loud and clear is the appeal to 'great' episodes in the history of science, and these episodes are told in respect to eventual falsifiability of theories which at first may have seemed like occult conjectures. One need not surrender the demaractionist ideal of falsifiability in phases of theory-deveopment where one can defend only by admittedly unsafe common-sensical arguments the scientific spirit of a research programme. In other words, in order to assume a falsificationist methodological posture, one need not hold only falsifiable theories.[96]

iv.

The heading of the section in which Lakatos introduces the notion of 'positive heuristic' reads:

> *Positive Heuristic: the construction of the 'protective belt' and the relative autonomy of theoretical science*[97]

The catchphrase here is "the relative autonomy of theoretical science". Lakatos appears to suggest that the introduction of a protective belt not only inaugurates a research programme but also creates the conditions which socially insulate it. On this account, the protective belt may originate a scientific discipline both as a social institution and an intellectual concern. However, it seems that to Lakatos the issue of the autonomy of science extends only to the question where and by whom scientific problems are generated[98]. But Lakatos's proposal can be extended to explain the insulation of scientific discourse from ideological discourse in general and thus to an explanation of phenomenon (Ph5):

> (Ph5) Though the theory of evolution by natural selection as well as some of its predecessors (theories of evolution and Malthus's principle of population) were subject to heated ideological and theological debate, Darwin's own writings as well as his exchange with many contemporary critics remained remarkably untouched by this controversy.

(Ph5) is closely related to (Ph4). The acceptance of evolution was a direct outflow from the shift of debate to the principle of natural selection. That shift of debate created a realm of discourse which presupposed that the process of evolution needs accounting for. And by the same token, that shift made 'evolution' a matter of purely theoretical concern in its relation to the explanatory principle of natural selection. The account for (Ph4) can thus be extended to an account of (Ph5) if one only provides a crucial 'missing link', namely, a theory of why the emergence of purely theoretical concerns should matter enough to the scientific community as to produce the pronounced watershed between theoretical and ideological concerns that has become manifest in (Ph5).

In the sociology and social history of science (Ph5) has been a most fruitful anomaly to historical-materialist and other interest-oriented approaches. It was fruitful insofar as in the course of accounting for (Ph5) from a Marxist

point of view, the importance of argumentation repertoires in science was first uncovered. However, that notion was not fully explored and the explanation of (Ph5) therefore remained unduly antagonistic to all internalist theories of science, whether philosophical or sociological.
Supplementing the interest-oriented approach and bringing it in tune with the Lakatosian explanation of (Ph4) amounts to delineating the 'professionalist argumentation repertoire'. And just as the 'inductivist' and 'demarcationist' repertoires were culled from philosophical theories of science, role and content of this latter repertoire can be culled from sociologist Robert Merton's theory on the normative structure of science.

In a review of Darwin-historiography, Steven Shapin and Barry Barnes criticize the ways in which historians have dealt with social Darwinism in general, and (Ph5) in particular.

> The central question seems to have been whether the author of the *Origin of Species* of 1859 was in any way either a "part of," or "responsible for," social Darwinism. The generally agreed-upon answer seems to be that the accused was innocent on both counts. This chapter [...] reflects upon the likely reasons for, and consequences of, such a protracted and expensive litigation.[99]

As they scrutinize the methodological presuppositions of the historians under review, Shapin and Barnes encounter the work of Robert Young who has come closest to the desired explanation of (Ph5). In the tradition of Marxist historiography, Young focusses on the ideological background of the nineteenth-century debate on "Man's Place in Nature". Within the context of his inquiry, (Ph5) clearly presents an anomaly:

> Lyell and Darwin [...] are, relatively speaking, the purest of scientists in the Victorian debate and as such are nearer to the position of physicists, chemists, and mathematicians.[100]

Indeed, the relative ideological purity of Darwin's writings appears to invalidate the Marxist explanatory model according to which science is (merely) a device for the construction of an ideological supersturcture which serves to justify the economic base of contemporary society.

> It can, of course, be argued that the base-superstructure distinction is only useful to historians of science as a vehicle for freeing themselves from the restrictions of the internalist-externalist distinction and that having once freed themselves, they should lay aside the model which has led so disastrously to economic reductionism in vulgar marxism. [...] My own current position is that the employment of a sufficiently subtle theory of mediations and interactions between socio-economic factors and intellectual life would make the base-superstructure model servicable once again, at least in an interim way, until we can develop a fully relational, totalising approach.[101]

Young founds his substitution of "vulgar marxism" by a "subtle theory of mediations and interactions" on Marx himself who in *The German Ideology* prepares the ground for an explanation of (Ph5).

> each new class which puts itself in the place of one ruling before it, is compelled, merely in order to carry through its aims, to represent its interest as the common interest of all the members of society, that is, expressed in ideal form: it has to give its ideas the form of universality, and represent them as the only rational, universally valid ones.[102]

Not discouraged by the methodological problem of how to empirically establish the efficacy of class interest within a "sufficiently subtle" historiographical approach, Young proceeds to evaluate Darwin in the terms set by Marx:

> It should be granted that the work and influence of Lyell and Darwin were less intentionally and obviously an expression of more basic socio-economic forces and structures than, for example, the work and influence of Chambers, Spencer, and Wallace. Similarly, their greater scientific prestige meant that those who employed their theories for socio-political purposes could claim a sounder foundation in the nature of things - in scientific laws - for their extrapolations and generalizations.[103]

The ideological concerns of Lyell and Darwin may not be reflected in their writings, but this makes them only particularly successful in the task of legitimizing those concerns. Phenomenon (Ph5) arises as scientists succeed in rendering the ruling ideas of the ruling class in "the form of universality [...] as the only rational, universally valid ones". (Ph5) now confirms rather than contradicts the base-superstructure approach: the fewer traces of their class-commitments scientists leave in their work, the more serviceable their work is going to be to their class. Along these lines, Young's more detailed analysis of Darwin's argument culminates in the claim that Darwin's success as a scientist consisted in assembling from a few familiar components a theory in the desired, putatively non-ideological form of universality: applying Lyell's uniformitarianism, Darwin simply extrapolated the model provided by artificial selection to the development of all biological organisms. Accordingly, Young concludes:

> Darwin's mechanism - in its nineteenth-century form and its nineteenth-century context - turned out to be a very frail reed, but in bending with the winds it allowed his real commitment to the uniformity of nature to contribute to the general movement of nineteenth-century naturalism. If we notice the extent to which the special status of natural selection was weakened by scientists, theologians, and philosophers, Darwin's achievement turns out to be much more like that of Lyell and of the other evolutionists: together, by a rather confused mixture of metaphysical, methodological, and scientific arguments which depended heavily on analogical and metaphorical expressions, they brought the earth, life, and man into the domain of natural laws.[104]

Young's explanation of (Ph5) leaves sociologists and social historians of science (like Shapin and Barnes) dissatisfied since it remains fundamentally ambiguous towards science and scientists in relation to overall society. As long as the problem is couched in terms of Darwin's intentions or "real commitments", Young's account invites speculations as to whether science or scientists actively promote ideological causes or whether the very objectivity of scientific results renders these results so malleable that society at large can put them to any ideological use at all[105]. What seems to be missing is a way of systematically assessing the relation of scientists and science to culture.

> For all that Young's work has been important and convincing, he has not succeeded in shifting the balance of historical attention away from the individual and his motives.[106]

And insofar as Young *does* attempt a shift "away from the individual and his motives", his statements

> represent a revealing gesture towards an institutionalized orientation in the history of science.[107]

That is, it may not be sufficient to shift attention from individuals to the institution of science as long as the institutional structure of science is thought to be differentiated from overall culture by a set of attitudes and motives shared by all scientists individually. If talk of individuals and motives does not help to clarify the scope of scientists' contributions to the justification of ideology, talk in terms of the institution of science as a particular, maybe arbitrarily defined aggregate of individuals and motives will not help either. Instead, the very question whether the institution of science is somewhat distinct from overall culture has to be raised as an empirical issue.

> We think and act on the basis of the resources our culture provides; and thereby we may add to these resources, and provide additional materials upon which the thought and activity of others may

> be constructed. There is no reason for treating Darwin, Newton, or indeed any scientist, as an exception to this. [...] Nonetheless, a culture need not be seen as an homogeneous whole. How differentiated it is at any time, how many "sub-cultures" exist within it, how "insulated" they may be from other sub-cultures and from the wider culture, and to what extent any individual's writings connect with this or that part - all these questions cannot be decided by a priori conceptions of how science *must be*, but only by "going and looking". Individualistic accounts of science, which relate the truth of science to its purity, anticipate the answers to empirical questions.[108]

In the case of Darwin and his wider culture, Shapin and Barnes suggest that it was not sufficiently differentiated to warrant recourse to a juxtaposition of scientific culture (in which Darwin participated) and society overall (which was ideologically supported by social Darwinism).

> In fact, there is every reason to suspect that the British intellectual scene of approximately 1859 was comparatively undifferentiated, with cultural resources being taken from a common stock and put to a wide range of uses. Robert Young has convincingly documented this "common context" [...] To the extent that Young [...] prove[s] to be right about the context, we should stop making differentiations inappropriate to that context.[109]

Shapin and Barnes do not explain (Ph5), but they suggest that the phenomenon itself is couched in questionable categories, that it simply does not exist, presupposing as it does a somehow problematic contrast between ideological and scientific debates[110]. By thus diffusing rather than solving problem (Ph5), Shapin and Barnes have shouldered another apparent anomaly: why would so many historians (including Robert Young) misperceive the phenomenon and frame it in terms of 'science' and 'culture'? Their answer to this problem is simple and compelling: Young structures the problem in terms of two cultures for the same reason and equally mistakenly as Darwin himself did. Where Darwin justified his own actions on supposedly 'scientific' grounds even though scientific discourse at his time was continuous with ideological discourse, Young accounts for Darwin's actions in terms of arguments and motive-ascriptions which are modeled after the currently (still) presumed separation of institutionalized science and wider culture.

> As Mills (1940) says, vocabularies of belief and motivation are to be treated as shared public resources, intelligible by reference to the community possessing them, and what that community does with them. Thus, the imputation of motives by professional historians is to be understood by *their* communal concerns and modes of practice. And, analogously, Darwin's self imputation of motives can only be considered in relation to communal concerns and practices in the context wherein they were uttered. Darwin's expressed motives do not provide a pipeline to his soul.[111]

Thus, without the reference to Marx and on methodological grounds different than his, Shapin and Barnes preserve Young's effort to provide a partial explanation of (Ph5) by bringing to the fore the common context of naturalists' debates in Victorian England. Their ensueing criticism is well-taken as they point out Young's failure to question his own as well as contemporary ascriptions of motive in terms of (ideologically) acceptable vocabularies of motive. However, from here on their argument becomes more questionable. Had Young been aware of ideological vocabularies of motive - it continues - he would have seen that only an undifferentiated wider cultural context provided the referent to Darwin's vocabulary, and he would have consequently realized that phenomenon (Ph5) was framed in a way which in turn is merely an artifact of contemporary society and *its* vocabulary of motive.

The two parts of Shapin and Barnes's argument against Young employ Mills's notion of vocabularies of motive in two distinct ways. Firstly, it is used to illuminate the dependency of any and all motive ascriptions on the vocabulary available to society at any given time. In this capacity, it compellingly questions Young's attempt at uncovering the 'real', 'extrascientific' commitment of someone who argues within the institution of 'science'. Indeed, sociological and philosophical theories of science should avoid explanations of social actions in reference to either professed or real motives. But having discovered the (merely) ideological foundation of ascriptions of motive, Shapin and Barnes employ Mills's notion secondly in order to altogether discount the social efficacy of vocabularies of motive. Used in this capacity, it leads to the counter-intuitive result that (Ph5) is not properly framed and that there is no clearly discernible tradition of science as a considerably autonomous social institution for the production of objective or, if one prefers, objectified knowledge. Surely, there is no way of convincing Shapin and Barnes that the contemporary view which clearly identifies such a tradition is more than false consciousness[112]. It *can* be shown, however, that Mills's proposal for a research-programme on vocabularies of motive provides for both, an ideological critique of ascriptions of motive *and* an account of the social efficacy of these vocabularies, i.e. that Mills's programme need not culminate in the denial of social institutions. On the contrary, social institutions are each defined by a vocabulary, and each is therefore able to perform very specific tasks such as 'scientific' negotiation towards the production of objective or objectified knowledge. In other words, one need not deny the ideological dependency of scientific thought (e.g. in the formation of basic concepts or in scientists' pursuit of goals), if one wants to elucidate the internal social structure of science which is defined by a self-imposed confinement to scientifically acceptable vocabularies of motive or argumentation repertoires. Having said this, one can further argue against Shapin and Barnes that the more comprehensive view on vocabularies of motive leads to a more comprehensive and therefore better explanation of (Ph5).

These contentions can be supported by first taking a brief look at three passages from Mills's stimulating and intriguing 1940 paper on *Situated Actions and Vocabularies of Motive*[113]. Just like Shapin and Barnes, he makes a compelling case against using reference to either professed or real or underlying motives in the explanation of social action.

> "Real attitudes" versus "mere verbalization" or "opinion" implies that at best we only infer from his language what "really" is the individual's attitude or motive. Now what *could we possibly* so infer? [...] All we can infer and empirically check is another verbalization of the agent's which which we believe was orienting and controlling behavior at the time the act was performed. The only social item that can "lie deeper" are other lingual forms. [...] What is reason for one man is rationalization for another. The variable is the accepted vocabulary of motives, the ultimates of discourse, of each man's dominant group about whose opinion he cares.[114]

Here, Mills uses the term 'ultimate of discourse' for 'vocabulary of motive'. Where Shapin and Barnes seem to think that these vocabularies will be used (merely) for *ex post facto* justifications of social actions which are performed independently of how they will be justified, Mills emphasizes the efficacy of vocabularies as ultimates of discourse when it comes to anticipating and evaluating the consequences of contemplated action.

> Motives are accepted justifications for present, future, or past programs or acts. To term them justification is *not* to deny their efficacy. Often

> anticipations of acceptable justifications will control conduct. ("If I did this, what could I say? What would they say?") [...] A man may begin an act for one motive. In the course of it, he may adopt an ancillary motive. This does not mean that the second apologetic motive is inefficacious. [...] It may strengthen the act of the actor. It may win new allies for his act. [...] When an agent vocalizes or imputes motives, he is not trying to *describe* his experienced social action. He is not merely stating "reasons". He is influencing others and himself.[115]

Social settings at once make available, constrain, and validate reasoning from a particular vocabulary. Rationalization is not 'mere' but legitimate rationalization - to understand science is therefore not simply to see that scientists can use presumably objectifying vocabularies to opportunistically embellish the choice of a given goal, but also to understand the social setting of science which legitimates and prescribes the employment of these particular vocabularies[116]. The question has now been raised just how a social institution legitimizes the employment of certain vocabularies. But it is rather obvious that e.g. the inductivist and demarcationist argumentation repertoires in science are legitimate at least for the somewhat banale reason that they can be successfully employed: they are good at what they do (i.e. they *do* discern rational grounds for belief, they *do* ensure the empirical content of theories, etc.)[117]. The availability, legitimacy, and validity of a certain limited number of repertoires marks a given social setting as a normative setting. The recognition of vocabularies of motive as ultimates of discourse, and the constraints on these vocabularies as indices of the normative character of social institutions, readmits the notion of vocabularies of motive to the task of explaining social action and phenomena like (Ph5): instead of explaining a particular action in reference directly to stated or inferred motives, one can explain in reference to available vocabularies of motive the patterned ways of framing social action in accordance with the norms of a institutional settings. Accordingly, Mills states the perspective for further research in the following terms:

> Rather than interpreting actions and language as external manifestations of subjective and deeper lying elements in individuals, the research task is the locating of particular types of action within typal frames of normative actions and socially situated clusters of motive. There is no explanatory value in subsuming various vocabularies of motives under some terminology or list. Such procedure merely confuses the task of explaining specific cases. The languages of situations as given must be considered a valuable portion of the data to be interpreted and related to their conditions. To simplify these vocabularies of motive into a socially abstracted terminology is to destroy the legitimate use of motive in the explanation of social actions.[118]

Mills's recognition of the (normative) efficacy of vocabularies of motive results implictly from his application of the "Thomas theorem" in the sociology of knowledge. Named after W.I. Thomas, it is a fundamental tenet in Robert Merton's sociological work, including, of course, his work in the sociology of science:

> "If men define situations as real, they are real in their consequences," wrote Professor Thomas. [...] The first part of the theorem provides an unceasing reminder that men respond not only to objective features of a situation, but also, and at times primarily, to the meaning this situation has for them. And once they have assigned some meaning to the situation, their consequent behavior and some of the consequences of that behavior are determined by the ascribed meaning.[119]

Indeed, it appears that much of the controversy on the existence and efficacy of institutional norms hinges on whether and to what extent one wants to take into the account the Thomas theorem when considering the acceptable vocabularies of motive. The controversy on the normative structure of science provides a case in point. In one of the seminal papers for the sociology of science[120], Robert Merton identified four norms of science as they

> can be inferred from the moral consensus of scientists as expressed in use and wont, in countless writings on the scientific spirit and in moral indignation directed toward contraventions of the ethos.[121]

The four norms are *universalism* which regulates equal access to the institution of science, *communism* which prescribes the shared ownership by the scientific community of scientific results, *disinterestedness* which delimits the system of goals and rewards that are legitimately available to scientists, and *organized scepticism* which mandates - as will be shown in more detail – that scientists adhere to the acceptable argumentation repertoires of science and that by implication they do not invoke societal argumentation repertoires as e.g. those of religious or ideological dogma. Since Merton's brand of sociology is firmly based on the Thomas theorem, it is a secondary issue whether these norms are instituted 'only' as acceptable and effective vocabularies of motive or in addition fortified by severe and stringently enforced sanctions. Indeed, Merton's later work explores these issues by focussing on scientists' ambivalence towards these norms[122]. And Michael Mulkay began to explore them by making explicit what is at least one manifestation of these norms.

> they are undoubtedly relatively standardized verbal formulations which are used by participants to describe the actions of scientists, to assess or evaluate such actions and to prescribe acceptable or permissible kinds of social action. [...] the standardized verbal formulations to be found in the scientific community provide a repertoire which can be used flexibly to categorise professional actions differently in various social contexts and, presumably, in accordance with varying social interests.[123]

With that last remark, Mulkay opens an intriguing research perspective which he formulates in questions such as

> How do scientists use vocabularies of justification inside their professional community? Are different vocabularies used in different social contexts; for example, in public as opposed to private media of communication? Are the more powerful groups and individuals better able to employ these vocabularies to serve their interests?[124]

Unfortunately, Mulkay does not rely much on the Thomas theorem and is consequently rather unclear on the efficacy of the normative vocabulary. This leads him to juxtapose the idea of normative structure and (mere) ideology in a way which - from the point of view of the Thomas theorem - appears rather curious, if not entirely untenable:

> I wish to suggest, therefore, that scientists have tended to select from their repertoire of accounts, those formulations originally taken by the functionalist interpreters to be the central norms of science; and that this version was selected because it served the social interests of scientists. It follows that the original functional analysis did identify a genuine social reality, but one better conceived as an ideology than as a normative structure.[125]

In the spirit of Thomas, Mills and Merton, Mulkay's abandoned project shall now be continued at least to the point where the content of the argumentation repertoire which contains the professional norms of the scientific community is partially elaborated. Both the inductivist and demarcationist repertoires were introduced as presenting options to practicing scientists who would select from them those arguments which would support their choices among alternatives in science. Though these repertoires may well be founded upon normative theories of rationality, their use is fairly unconstrained[126]. The professionalist repertoire is also open to opportunistically selective employment. However, in contrast to the inductivist and demarcationist repertoires, appeal to the professionalist repertoire has normative character. For, only as the institution of science is socially defined and identified by its self-imposed confinement to certain argumentation repertoires (including, of course, the inductivist and demarcationist repertoires), that adherence is enforcable as a scientific norm[127]. Only then, transgressions of available repertoires become violations of professionally acceptable conduct, and appeal to the professionalist repertoire can be employed to publicly identify them as such. A famous and dramatic exchange at the meeting of the British Association for the Advancement of Science at Oxford in 1860 illustrates this point.

> It will be remembered that the Bishop [Wilberforce of Oxford], towards the end of a long speech where he denounced the Darwinian doctrines, had turned to Huxley and mockingly asked him whether he reckoned his descent from an ape on his grandfather's or his grandmother's side? - to which Huxley retorted: "If the question is put to me, would I rather have a miserable ape for a grandfather or a man highly endowed by nature and possessing great means and influence, and yet who employs those faculties and that influence for the mere purpose of introducing ridicule into a grave scientific discussion - I unhesitatingly affirm my preference for the ape."[128]

One might first try to interpret the Bishop's remark as quite innocuous: using the rhetorical equivalent of a *reductio ad absurdum*, the Bishop sceptically assessed a deductive consequence or extrapolation of Darwin's theory, namely an implication by the theory of common descent.

However, no available argumentation repertoire provides for this sort of argument and Huxley's response makes quite clear that this violates the norm of *organized* scepticism. The Bishop's argument does not organize the technical, verbal or communal resources of the scientific community. That is, in support of his sceptical contention Bishop Wilberforce did not establish technical (test-)conditions, he did not adhere to the available argumentation repertoires, and he did not solicit the critical scrutiny of other scientists. Instead, he spoke only with the force of personal authority which does not by itself count for much in the context of scientific negotiation. Thus, Huxley's response articulated a communal concern with the enforcement of the norms governing scientific conduct. A contemporary account of the episode accordingly interprets the Bishop's remark as a violation and Huxley's as a defence of *organized scepticism*.

> It was sad, indeed, to think that the opponents of the theory [...] summon[ed] to their aid a species of oratory which could deem it an argument to ask a professor if he should object to discover that he had been developed out of an ape. The professor aptly replied to his assailant by remarking, that man's remote descent from an ape was not so degrading to his dignity as the employment of oratorical powers to misguide the multitude by throwing ridicule upon a scientific discussion. The retort was so justly deserved, and so inimitable in its manner, that no one who was present can ever forget the impression it made. Happy are we to be able to escape from such recriminations, for there is some chance of a satisfactory solution when we can appeal to physical principles.[129]

The norm of organized scepticism is used to enforced adherence to the available argumentations repertoires. As such, it functions like the exclusionary clause in a recursive definition, namely a definition of (contemporary) 'scientific activity'[130]:

> 'Scientific activity' is (socially) defined as public negotiation on alternatives in science in such a way that
> (1) arguments from the inductivist repertoire are permissible,
> (2) arguments from the demarcationist repertoire are permissible,
> (...)
> (n) arguments from repertoires other than those listed in (1) through (n-1) are not permissible.
>
> (n) is enforced by employment of the professionalist repertoire with appeal to the norm of organized scepticism.

The professionalist repertoire (just like the inductivist and demarcationist repertoires) makes use of organized scepticism in a flexible manner. It is allows for more than the mechanical employment of a rule or a set of rules. Instead, there are patterns of lenience and strictness in the enforcement of adherence to acceptable arguments. A variable in such patterns is e.g. a scientist's past performance. Darwin had scrupulously established himself as a diligent observer and cautious theorist before advancing his revolutionary theory. Indeed, a few years after publication of the *Origin*, Darwin gave the following advice to a young scientist:

> I would suggest to you the advantage at present,

> of being very sparing in introducing theory in your papers (I formerly erred much in Geology in that way): *let theory guide your observations*, but till your reputation is well-established be sparing in publishing theory. It makes persons doubt your observations.[131]

This passage suggests that a well-reputed scientists have less to fear as they stress the limits of currently acceptable argumentation. But even Darwin who had already built himself a reputation continued to deliberately cultivate it after publication of the *Origin*:

> Anything said by myself in defence would have no weight; it is best to be defended by others or not at all.[132]

Darwin's attentiveness to his professional standing paid off: many of his critics gave weight to his arguments by citing his impeccable scientific stature[133].

The other norms of science play a role similar to that of 'organized scepticism'. In some way or other, they all are normative extensions of methodological argumentation repertoires. It seems that 'univeralism' plays hardly any role when it comes to evaluating a new theory like the theory of 'natural selection'. The merits of equal access to domains of thought which were previously imbedded in authoritarian structures were more pronounced in the 17th than the 19th century. The claim that Darwin transplanted the species-question from a hierarchically structured community (church and theology) into the domain of egalitarian science, would not be fair to most of Darwin's critics[134]. The norm of disinterestedness is also somewhat marginal in the case at hand. Disinterestedness is related to organized scepticism in analogy to the relation between goals and arguments in science: it regulates the scope of publicly statable goals, i.e. it fends off scientific cynicism and compensates for the incurred loss of ego-satisfction. A cynical attitude might employ admissable arguments in an openly strategic manner, e.g. in stating that such and such is the intended goal but that in regard to acceptable repertoires the following manner of argumentation will have to be adopted. The norm of disinterestedness thus contributes to a definition of the scientific persona. Also, arguments in respect to disinterestedness play a central role in the adjudication of priority-conflicts: the otherwise disinterested scientists deserves at least recognition where recognition is due.

Finally, the norm of communism enters into an explanation of (Ph5). It is also employed in a variety of ways. For instance, Darwin notes that the general availablity of a new theory to a given scientific community can be appreciated only by that community itself:

> I believe that Hopkins is so much opposed because his course of study has never led him to reflect much on such subjects as geographical distribution, classifiction, homologies, &c., so that he does not feel it a relief to have some kind of explanation.[135]

This is, so to speak, a negative and restrictively qualifying use of communism as a norm: shared access to a new theory does not mean the same to persons at different locations within the scientific community. A positive use of that norm is manifest in the countless assertions that Darwin's theory has founded a science, that is, that it has put to work in a productive manner and in a variety of areas the resources of science.

> Whatever opinion may be entertained of the speculations on the origin of species sketched out by Mr. Darwin [...] there is no doubt that in its entirety his theory is one which for many yers to come must receive the earnest attention of the scientific world; for whether the law of the necessity of organic variation and development as dependent on external circumstances attendant on the general "struggle for life" be universal in application or not, Mr. Darwin has at any rate opened out a new vein of reflection and investigation which must be followed out until the new theory be either disproved or proved from its first causes to its final results.[136]

Reflections on science by scientists are sometimes read as philosophical statements. But instead of establishing methodological doctrines, these utterances more often than not employ the professionalist repertoire in order to defend or attack a given choice among alternatives in science. The professionalist repertoire is a vital component of an elaborate process which transforms more or less timely ideas into theoretical form. By subjecting a theory (maybe, any theory at all) to a sustained treatment with scientifically acceptable arguments, that theory will ultimately emerge as a properly scientific theory[137].

Notes to Chapter 4

[1] See above, chapter 1, pp. 32 and 39.

[2] This (L*) is a proper specification of the (anti-theoretical) law (L*) which was discussed and tentatively adopted in the first chapter (see above, p. 18): the latter like the former rests on a presumption of rationality in scientific argumentation.

[3] Two features shared by many theories of science may be accountable for taking this limiting case as paradigmatic: i) the attribution of an essentially deductive structure to scientists' belief-systems; ii) the assumption that the deductive structure is a complete representation of these belief-systems and that e.g. *modus tollens* operates quasi-automatically within this structure. Kyburg [1983], pp. 300f gives a concise characterization of this jointly attributed deductive structure:

> *First:* There is a deductively closed and strictly consistent set of statements accepted by the scientist. This set may be empty (for a superextreme Bayesian); or it may include any statements that in a given context are regarded as "unproblematic"; or it may even include the very general sorts of statements that characterize a paradigm or a research program.
>
> *Second:* There is a set of statements (axioms of theories, hypotheses) that are considered or contemplated. These are the statements that are up for test. They may be quite general, or they may be quite specific, but in any case they have empirical content, and they are to be confirmed, in some sense or other, by their instances or consequences, or they are to be refuted by counterinstances.
>
> *Third:* There is a set of statements that is directly testable, statements that express the results of experimentation or observation directly and incorrigibly.
>
> According to this general view, scientific procedure consists in putting statements of this third class to experimental verification or refutation; the result may be the deductive refutation of a statement belonging to the second class (Popper); or the increase in the degree of confirmation of the hypothesis in the second class

> (Carnap); or the general rearrangement of degrees of belief assigned to items in the second class (subjectivist Bayesian); or the verification of a statement in the second class, in virtue of general assumptions embodied in the first class (Lakatos, Sneed); or the basis for a computation providing confirming instances of statements in the second class (Glymour); or it might be the grounds for the rejection of a null hypothesis (for the practicing statistician).

Kyburg goes on to refute each of the three basic assumptions. Here they are all accepted in some version or another, if only insofar as they provide a powerful representational device. It is proposed, however, that the general assumptions in the first class allow at best occasionally a unique and unequivocal verdict on the statements of the second class in relation to the statements in the third class. However, the overall weight of arguments generated with the help of the basic assumption in the first class (above and beyond those arguments provided by the statements in the third class) can effect a great deal towards ultimate persuasion and consensus-formation. Compare Darwin's letter to G.H.K. Thwaites (March 21st 1860) in Darwin [1903], Vol. I, p. 145 (letter # 97):

> As for changing at once one's opinion, I would not value the opinion of a manwho could do so; it must be a slow process.

A few more remarks on the problem of consensus-formation can be found in the conclusion.

4 The inclusion of 'type of argument' permits a very liberal reading of (Ph1) as it allows to consider all references to empirical evidence as one type of argument. (Ph1) asserts that one will often find other types arguments besides this most prominent one, concerning e.g. the beauty and simplicity, usefulness or timeliness of the theory.

5 This phenomenon cannot be established here. It will be exemplified by the remainder of this chapter in the single case of Darwin's argument for 'natural selection'.

6 Again, the phenomenon cannot be established here. For the case of 'natural selection' consider for instance the following concise criticism of Darwin's and Wallace's theory by Samuel Haughton (quoted in Hull [1983], p. 216):

> If it means what it says, it is a truism; if it means anything more, it is contrary to fact. Q.E.D.

And for a more puzzling formulation, see Picted [1983], p. 146:

> Thus we find ourselves in a singular position. We are presented with a theory which on the one hand seems to be impossible because it is inconsistent with the observed facts and on the other hand appears to be the best explanation of how organized beings have been developed in the epochs previous to ours.

7 These qualifications and provisos should not be taken to imply that the notion of 'argumentation repertoire' is empirically inaccessible or insignificant. They imply only that it takes considerable scrutiny to discern and establish relevant evidence in order to establish the efficacy of any argumentation repertoire. Two kinds of relevant evidence come to mind. *Firstly*, complaints about and controls over deviant scientific discourse constitute such evidence. For instance, Darwin's rare departures from sober scientifically acceptable argumentation did not escape notice by his critics. Wollaston [1983], p. 129 paraphrases and quotes Darwin and then rejoinds, making use of one of the repertoires which will be discussed below:

> There is one point, however, according to Mr. Darwin's own confession, which has struck him much: viz. That all those persons who have most closely investigated particular groups of animals and plants, with whom he has ever conversed, or whose treatises he has read, are firmly convinced that each of the well-marked forms was at the first independently created. But, says he, the explanation for this is simple: from long-continued study they are thoroughly impressed with the distinctions between the several races, and they ignore all general arguments, -refusing "to sum up in their minds slight differences accumulated during many successive generations." But is this more, we may ask, than special pleading? If anybody is capable of forming an opinion on the origin of species, it surely must be those who have most closely studied them

Wollaston, a rather benevolent critic, proceeds by substituting Darwin's flawed argument for a better one which serves the same end.

> The true explanation seems to be this: not that the study of small details unfits an observer for wider areas of thought, but simply that a generalizing mind is of a higher stamp, and therefore less common, than one of an opposite tendency; so that there are more *collectors* in the world than generalizers. (p. 129f.)

For more polemical passages on the same transgression see Owen [1983], pp. 181f. and 210. And for similar charges in in defence of Darwin see e.g. Hutton [1983], p. 293. Controls on deviant discourse are institutionalized in the refereeing system of scientific journals, see Merton and Zuckerman [1973]. *A second type of evidence* for the detection of argumentation repertoires can be construed by taking sets of thematically correlated formal and informal statements as single pieces of evidence, where one set or piece of evidence would include e.g. informal utterances in the lab, scribbled notes to oneself, letters, statements at conferences and in discussions, and finally the formulation printed in a journal - all on the same subject-matter. The distribution of, say, ad-hominem remarks in relation to arguments from evidence will be be gradated within each set as one moves from less formal to more formal communications. And this affords clues as to the constraints governing formal settings. Knorr-Cetina [1981], pp. 94-135 introduces and discusses an example of this sort of evidence at length. However, her very rich and fascinating example is used to illuminate the constraints governing public and formal discourse only in one respect. She shows how the published version of a paper does not tell the story of its genesis: the "scientific paper hides more than it tells on its tame and civilised surface" (p. 94), it decontextualises and is "marked by a consipicuous *avoidance* of arguments" (p. 113). Yet, it should be equally interesting to find out what story the constrained published paper *does* tell, and why one would want to tell this story rather than the suppressed version. (For more on the argument that Knorr-Cetina's approach needs to be supplemented in some such way, see above page 48, note 40 to chapter 1.) In Darwin's case, one would notice e.g. that the word 'materialism' does not once occur in the first edition of the *Origin* while it figures prominently in the notebooks (see Gruber [1981], esp. pp. 38-42 and 201-217). Indeed, the notebooks yield explicit testimony to Darwin's resolve to avoid that word in order to diffuse not only ideological and theological objections, but primarily objections concerning the purely scientific standing of his argument. Compare Notebook M, p. 57 (Darwin [1980], p. 16):

> To avoid stating how far, I believe, in Materialism, say only that emotions, instincts degrees of talent, which are hereditary are so because brain of child resembles parent stock.

It should be noted that speculations concerning Darwin's 'delay' in publishing his theory, often point to this effort at distilling a formulation that most nearly contains nothing but scientifically acceptable arguments, compare for instance Richards [1983].

8 This discussion cannot be summarized here. It may suffice to list two works which evaluate the possibility of discerning such laws quite differently. Zilsel [1941] is exceedingly optimistic, while Danto [1965] hesitates carefully.

9 This formulation of Hume's problem of induction has been taken from Stegmüller [1975], p. 17.

10 It will be remembered that Peirce's threefold distinction was preserved in the characterization of 'scientific activity' in chapter 2. It was modified only to accomodate various kinds of inductive acceptance at various stages of deductive exploration (compare the flow-chart on p. 127).

11 'Retroduction' is another term used by Peirce for 'abduction'.

12 Peirce [1958-60], 6.475.

13 Peirce [1958-60], 5.375.

14 It has been maintained that Popper's notion of 'corroboration' designates a form of (inductive) justification in disguise. There will be a brief discussion of this (mistaken) impression in the upcoming section on the demarcationist argumentation repertoire.

15 This is nicely exemplified by the index of Popper [1968], 469f. The entry "Belief" refers the reader to "Conviction" and to "Belief, 'rational' degree of". Under "Conviction" one finds "Conviction, feelings of, irrelevant to scientific discussion", followed by 9 page-references. – Compare Popper [1975], p. 25:

> I do not believe in belief.

16 Popper [1975], pp. 1-31.

17 Popper [1975], p. 4.

18 Popper [1976], p. 145.

19 Popper [1975], p. 6. Popper does not attempt to justify this principle. He merely states that it is

> admittedly a somewhat daring conjecture in the pschology of cognition or of thought processes (p. 6)

Clearly, as long as it has not been justified, H_{Ps} has not been solved on Popper's strategy and the formation of belief is not accounted for. However, given the curious 'solution' to the logical problem of induction, it is not at all clear whether Popper actually employs his principle of transference - especially since towards the end of his article he takes a completely different, much more plausible, and far more limited approach towards the solution of H_{PS}.

20 Popper [1975], p. 7.

21 Ibid.

22 Ibid.

23 Ibid.

24 The claim that Popper's answer to L_2 addresses Hume's logical problem rests on a fallacy which cannot have escaped his attention. It presupposes that Popper is engaged in a chain of reasoning which invites carricature of the following kind:

(1) L_1: Do triangles have four sides?
(2) L_2: Do triangles have four sides or do bicycles have two wheels?
(3) Bicycles have two wheels: L_2 is answered.
(4) L_1 is answered since L_2 (a generalization of L_1) is answered.

Popper's trick-argument (wittily employed as such) draws philosophical attention to an epistemological problem which will be further discussed in the next section: how *does* the answer to L_2 relate to the problem of induction - if it has not solved Hume's problem, has it rendered it obsolete?

25 And indeed, later one in Popper [1975], p. 26 one finds the assertion:

> I do not regard the psychological problem of induction as part of my own (objectivist) theory of knowledge.

26 Popper [1975], p. 27.

27 Reichenbach [1961], p. 348:

> Is it necessary, for the justification of inductive inference, to show that its conclusion is true? [...] The proof of the truth of the conclusion is only a sufficient condition for the justification of induction, not a necessary condition.

28 Reichenbach [1961], pp. 350f.

[29] Reichenbach [1961], p. 350.

[30] Reichenbach [1961], p. 340.

[31] Reichenbach [1961], p. 355.

[32] Reichenbach [1961], p. 356.

[33] Reichenbach [1961], pp. 351f.

[34] Reichenbach [1961], p. 353. - Some analogue to this "method of anticipation" is needed to supplement Popper's suggestion that belief is pragmatically justified in virtue of one's choice of the best method. Indeed, Popper's frequent use of the word "guess" corresponds rather neatly to Reichenbach's "blind posit".

[35] To actively anticipate revision of the best available data is not belief, but suspension of belief. At the same time one can believe that a given value represents the best data available right now. While believing a theory (i.e. believing it to be true) is to be ready to act upon its truth, believing that some value is presently the best available value implies only that one will keep looking for better values.

[36] See Levi [1967], pp. 35f. for a definition of 'ultimate partition'.

[37] Stegmüller [1975], pp. 59ff. subjects Carnap's earlier work to criticism very similar to the one raised here against Popper and Reichenbach. He concludes (p. 62):

> It seems to me that this problem - like all theoretical variations on Hume's problem of induction - is unsolvable as long as one follows Carnap's original sense of purpose and interprets his theory as a theory on the confirmation of hypotheses. (Solange man Carnaps Theorie als eine Theorie der Bestätigung von Hypotheses interpretiert, also seinem ursprünglichen Selbstverständnis folgt, ist dieses Problem, so scheint es mir, unlösbar - *wie alle theoretischen Varianten des Hume-Problems unlösbar sind.*)

The later, largely unpublished Carnap is to have considered inductivism as a search for norms governing decisions under risk. Since (according to Stegmüller) one cannot bet on the truth of natural laws, it is a desirable result of this decision-theoretic turn that no probability measures greater than 0 can be established for natural laws (p. 69).

38 For a more complete and infinitely more detailed account see Levi [1980]. His somewhat ideosyncratic version of inductivism is presented here.

39 Levi [1979], p. 421.

40 The ensueing presentation again follows Levi [1980]. Having solved this problem, he refers to his position as 'infallibilist' and 'corrigibilist' at the same time (see e.g. his [1983]).

41 Levi [1980], pp. 34-40.

42 As Lakatos [1978], p. 104 points out, all philosophies of science have their "characteristic victorious" examples from the history of science. As of yet, this particular inductivist methodology has no well-established paradigm example to point to. However, Rudwick [1974] unwittinlgy provides one: according to his conclusion on pp. 177f., his case-study was originally conceived in reference to Lakatos and Feyerabend rather than the inductivist methodology it illustrates so well.

43 Darwin [1980b]. - The following story is told in fascinating detail by Rudwick [1974].

44 Engelhardt and Zimmermann [1982], pp. 349-368 provide a more extensive interpretation of actualism as a 'regulative principle'.

45 Darwin [1903], Vol. 2, p. 188 [# 524].

46 Darwin [1887/1893], Vols. 1, pp. 68f./57. See also Darwin [1958], p. 84.

47 Rudwick [1974], p. 169. - The "alternatives, C... etc." should be represented as a 'residual hypothesis' in the ultimate partition: the residual hypothesis states as much as "none of the above". In many cases, only the inclusion of a residual hypothesis will supplement the exclusive alternatives so that an exhaustive set of potential answers is formed that properly constitutes an ultimate partition. Whether a residual hypothesis is needed or not is also subject to scientific negotiation.

48 Hull [1983], pp. 25f.

49 To be sure, it takes some idealization and simplification to divide the relevant problem-areas of the inductivist repertoire into four distinct segments.

For instance, no strict chronological order and no clear-cut systematic distinction is implied here between the specification of a demand for information and the choice of an information-ranking function. After all, a better explanation of the Parallel Roads was in demand only in respect to the deficiency (in informational terms) of extant explanations. Levi [1979], p. 421 indicates that informational value has to be fixed relative to demands for information and that therefore there can be no independent characterization of information-ranking functions, demands for information and ultimate partitions of potential answers. An information-ranking

> function is an expression of a demand for new information in a specific context of inquiry. [...] When demands for information become controversial, this conflict is in turn often a manifestation of a conflict between different regulative ideals each of which specifies directions for worthwhile research.

A similar difficulty arises when confirmational commitments are treated as relatively independent of bodies of belief: for instance, a firm belief that members of some race cannot perform certain intellectual feats will not be challenged by scientific arguments which were developed by members of that race - functioning as a confirmational commitment, the belief in question effectively devalues certain kinds of evidence relevant to itself.

50 Psychologically and as a matter of internal representation, 'preference among all available theoretical options' and 'belief' does not 'feel alike' at all. However, this feeling may well be due to the effects of centuries of belief-talk rather than to ascertainable conceptual differences: as long as 'preference' and 'belief' do not yield at least partially distinct consequences in action, they are (on pragmatist grounds) conceptually indiscernible. For the purposes at hand, no distinction needs to be made. Of course, there is the phenomenon of grudging or reluctant preference in the absence of *any* cogent hypothesis. Reluctant preference, however, is belief in the residual hypothesis of the ultimate partition, i.e. the true account is thought to be among the as of yet unknown accounts covered by that residual hypothesis. - Incidentally, in a domain which is held to deal with 'belief' in a deep, non-operational sense, theologian Hans Kung in his [1980] treats the question of belief in the existence of God as a matter of rational preference by eliminative induction among a given set of available options.

51 These questions are not pursued here any further. However, they will be raised again in the conclusion.

52 Darwin [1959/1962/1859/1981], Introduction, 21-22 and 26/pp. 26f./pp. 3f./pp. 66f. Huxley stated the demand the same matter far more bluntly in an essay on the reception of Darwin's theory which is included in Darwin [1887/1893], Vol. 2, p. 197/Vol. 1, pp. 550f.

> That which we were looking for, and could not find, was a hypothesis respecting the origin of known organic forms, which assumed the operation of no causes but such which could be proved to be actually at work. We wanted, not to pin our faith to that or any other speculation, but to get hold of clear and definite conceptions which could be brought face to face with facts and have their validity tested. The 'Origin' provided us with the working hypothesis we sought. Moreover, it did immense service for freeing us for ever from the dilemma - refuse to accept the creation hypothesis, and what have you to propose that can be accepted by any curious reasoner?

See also Huxley [1906], pp. 330f.

53 There was *no* rival to 'natural selection' in respect to the demand for information as specified by Darwin. That is, the theory of special creation was a rival of the theory of evolution and thus also of the theory of evolution by natural selection, but not of the theory of natural selection. The demand for the latter theory arose only to the evolutionist (see Huxley's statement in the previous note). See Wollaston [1983], p. 128 for the creationist's ultimate partition and an (on inductive principles) certainly rational way of making a selection from it:

> The opinion amongst naturalists that species were independently created, and have not been transmitted one from the other, has been hitherto so general that we might almost call it an axiom. True it is that we cannot prove this; but then, on the other hand, we cannot prove the converse; and, since of two unproveable propositions we have a right to take our choice, the former has been universally accepted, as most in accordance with the intelligible announcements of revelation

See also Pictet [1983], p. 146 (quoted above in note 6), and (criticizing Darwin's construal of the ultimate partition) Hopkins [1983], p. 268 and Owen [1983], p. 182ff., also Fawcett [1983], p. 279f.

[54] Letter (# 94) to J.D. Hooker (February 14th, 1860) in Darwin [1903], pp. 139f.

[55] Darwin [1959/1962/1859/1981], IV, 227/p. 114/p. 108/p. 153). See IV, 125/p. 105/p. 96/p. 142:

> as modern geology has almost banished such views as the excavation of a great valley by a single diluvial wave, so will natural selection, if it be a true principle, banish the belief of the continued creation of new organic beings, or of any great and sudden modification in their structure.

With the advent of 'punctuated equilibrium', the validity of Lyell's information-ranking function has been called into question.

[56] To J.D. Hooker (Dec. 24th 1858) in Darwin [1887/1893], Vol. 2, pp. 142f./Vol. 1, p. 498.

[57] Owen [1983], p. 209.

[58] Darwin [1887/1893], Vols. 2, pp. 241/36f.

[59] Phenomenon (Ph3) can now be more poignantly reformulated as:

> (Ph3) Darwin believes the theory of natural selection to be true without claiming to have proved or demonstrated its truth.

A cursory glance at the concordance to the *Origin* (Barrett et.al. [1981]) shows how very frequently Darwin in the *Origin* used the words 'believe', 'belief', and 'convinced'. On the other hand, 'prove', 'proof', 'demonstrate', and 'demonstration' occur only very rarely and often in a conditional ("if it could be proved") or negative ("incapable of demonstration") mood.

[60] Hull [1983] repeatedly makes the point that pervasive contemporary notions of proof and demonstration and their role in scientific inquiry are methodologically and historically mistaken. That is, they are integral part of an acceptable argumentation repertoire even though they are methodologically questionable. See for instance pp. 272ff. as he comments on Hopkins [1983].

[61] His principle of transference (see above, p. 193) only rules on when one is justified to believe that a theory is false.

[62] Examples for the rational adoption of what on Popperian analysis may turn out to be a mystery-raising

attitude, need not involve relatively far-out questions such as whether one can rationally hold certain religious beliefs. The passage by Lichtenberg which was quoted above on p. 60 provides a splendid example from the realm of scientific activity, torn as it is between abductive awe for a theory and weariness concerning its scientific status. - Compare also the reference to Küng [1980] in note 50 above.

63 Stegmüller [1975], p. 2:

> Zwischen den Theorien Poppers und Carnaps bestehen überhaupt keine Berührungspunkte. [...] Die Poppersche Bewährungstheorie betrifft die *theoretische Beurteilung von unverifizierbaren Hypothesen*. Die Carnapsche Theorie betrifft die *Aufstellung von Normen für menschliche Entscheidungen unter Risiko*.

"Absolutely no points of contact" refers to the two respective programmes in the philosophy of science. To be sure, that a given theory is scientific (as established within the demarcationist vocabulary) may enter as evidence into the (inductive) judgement concerning its acceptability. And conversely, all arguments showing that one can or should believe it to be true, make its construal as falsifiable and all attempts at falsification even more desirable.

64 The potential of at least rhetorical conflict between the two repertoires and corresponding postures is another instance of (Ph2), see above p. 187.

65 For Popper's own reflections on the cultural value of critical rationalism, see his [1966].

66 One can show here how the two repertoires interlock from the point of view of the inductivist repertoire. The demarcationist repertoire is needed to negotiate stronger minimal versions of information-ranking functions which make it more likely that informational value is attributed only to empirically significant hypotheses. Thus, the result that a theory is scientific on demarcationist grounds, becomes input into the inductive negotiation on the acceptance of that theory. Also, certain confirmational commitments will be negotiated within the demarcationist repertoire, especially the formulation of the conditions under which one will contract one's corpus of knowledge.

67 Popper [1976], pp. 79f.

68 Popper [1976], p. 104.

69 See e.g. Salmon [1968], pp. 25-29.

70 Stegmüller [1975], p. 10:

> Poppers Vorschlag lautet nun: Man soll in jedem erfahrungswissenschaftlichen Forschungsgebiet unter den miteinander konkurrienden Theorien die *empirisch gehaltvollste* auswählen – also diejenige *mit den meisten potentiellen Falsifikatoren*, d.h. die *riskanteste* oder *logisch unwahrscheinlichste* –, aber nicht etwa, um sie zu akzeptieren, sondern *um sie einer strengen Prüfung zu unterwerfen* (genauer: der strengsten Prüfung, die man sich ausdenken kann).

Compare also pp. 19ff. and 36f.

71 As pointed out before (see above, note 50), it is not possible to distinguish 'believing an account to be true' from 'preferring one account among all competing accounts'. Popper can maintain that there is a difference between 'preference' (non-justified) and 'belief' (presumably justified) only insofar as he makes somewhat counter-intuitive use of the term 'preference': the termd does not designate the mental experience resulting from the formation of a preference (and of the simultaneous adoption as true that the preferred theory is the best or true theory).

72 Popper [1968], p. 315.

73 Popper [1962], p. 218.

74 For instance, to the demarcationist repertoire belongs all negotiation on how to construe 'falsifiability' in respect to condition (D:) or to stronger condition (D*) (see above, pp. 80 and 101). Also, all negotiation on the admissability of given ad-hoc hypotheses belongs here. It involves the formulation of criteria for content-increasing versus content-decreasing ad-hoc hypotheses.

75 Popper [1968], pp. 54f. frames the problem of negotiating such questions within demarcationism and critically employed metaphysics:

> Methodological rules are thus closely connected both with other methodological rules and with our criterion of demarcation. But the connection is not a strictly deductive or logical one. It results, rather, from the fact that the rules are constructed with the aim of ensuring the applicability of our criterion of

> demarcation; thus their formulation and acceptance proceeds according to a practical rule of a higher type. [...] not a few doctrines which are metaphysical, and thus certainly philosophical, could be interpreted as typical hypostatizations of methodological rules.

Later in the *Logic of Scientific Discovery* (p. 207), Popper presents an example:

> the doctrine that *all* observable events must be explained as macro events [...] Like other doctrines of its kind, this seems to be a metaphysical hypostatization of a methodological rule which in itself is quite unobjectionable. I mean the rule that we should see whether we can simplify or generalize or unify our theories by employing explanatory hypotheses of the type mentioned

76 As a complement to note 66 above, one can now show from the demarcationist perspective where the inductivist and demarcationist repertoires interlock. The inductivist repertoire is needed to provide (tentative) beliefs which provide input (i) as critically employed metaphysics into the precise demarcationist (i.e. falsificationist) construal of parameters for the theoretical evaluation of rivaling scientific theories; (ii) as science-generative metaphysics (positive heuristic) into (falsifiable) scientific theories generated. Also (iii), (tentative) beliefs in certain empirical truths are required to establish the appropriate background knowledge that will render theories independently testable (see the discussion in chapter 2, esp. p. 116). Popper need not and does not deny that (metaphysical) beliefs play these roles in scientific development: he can still maintain that the specific virtues of scientific discourse lie in the relative insignificance of belief when it comes to formulating, presenting, and evaluating theories. Indeed, Popper [1976], p. 80 considers as a major advantage of his demarcation criterion that it does not render metaphysics nonsensical but only non-scientific:

> It was clear to me that all these people [i.e. members of the Vienna Circle] were looking for a criterion of demarcation not so much between science and pseudoscience as between science and metaphysics. [...] they were trying to find a a criterion which made metaphysics meaningless nonsense, sheer gibberish, and any such criterion was bound to lead into trouble, since metaphysical ideas are often the forerunners of scientific ones.

Schäfer [1974], pp. 81-99 discusses the various roles which Popper assigns to metaphysics in relation to science (i.e. its generative, critical, and explicative functions) and on pp. 64-73 also Popper's conventionalism concerning the acceptance of background knowledge and basic-propositions. Popper [1962], pp. 238-240 emphasizes the role of background knowledge. And the latter part of Radnitzky [1979] structures the problem-areas in which inductivism and demarcationism interlock from the demarcationist point of view.

77 Lakatos [1978b], p. 48.

78 Lakatos [1978b], pp. 50-52. For a general discussion of Lakatos's 'positive heuristics' see Urbach [1978]. In it Urbach suggests considerations which will enter into a negotiated assessment of heuristic power.

79 'Falsifiability in a significant sense' means 'falsifiability in respect to condition (D*)' as opposed to easily satisfiable conditions (C:) and (D:).

80 Michod [1981] discusses the role of positive heuristics in a later phase of evolutionary biology.

81 Romanes [1896], pp. 251f.

82 Ellegard [1958], p. 17 speaks of

> one of the paradoxes of the subsequent development of opinion: though it is practically certain that the evolution theory would not have been established at all if Darwin had not been able to support it by means of the naturalistic theory of Natural Selection, yet the majority of the general public, and a good many scientists, refused to accept the Natural Selection theory, while allowing themselves to be converted to evolutionism.

Darwin himself stated (Ph4) at least once in a letter to J.D. Hooker (July 28th 1868) Darwin [1903], Vol. 1, p. 304 (letter # 222):

> I am glad to hear that you are going to touch on the statement that the belief in Natural Selection is passing away. I do not believe that even the *Athenaeum* would pretend that the belief in the common descent of species is passing away, and this is the more important point. This now almost universal belief in the evolution (somehow) of species, I think may be fairly attributed in large part to the *Origin*.

83 Popper [1976], p. 171. - Note that the term 'metaphysical' designates the non-testability of the hard core. Its use is not supposed to imply that the propositions of that hard core say nothing meaningful about the world.

84 Popper [1976], p. 171. - This point is reiterated by other Popperian interpretations of evolutionary theory. Compare for instance Manser [1965], Barker [1969], and Lee [1969]. Manser's reflection on the concept of evolution contains the following telling remarks (p. 29):

> It is hard to understand the notion that it is a property of matter to develop into more complex forms and ultimately into livings things. More seriously, it is meaningless as a hypothesis because no experimental test could be devised for it, even in pinciple.

It is easy to agree with the first statement: it expresses the problem posed by disposition predicates in scientific explanation. However, the very difficulty of understanding such dispositional properties may in the due course of scientific negotiation yield a testable construal of that property. After all, the ways in which predicates designate properties are not immutable and therefore hard cores can be preserved through appropriate negotiation on the protective belt. It thus appears that critics like Manser and Barker conceive Popperian doctrine too narrowly. Popper's and Lakatos's protestations notwithstanding, for historiographical purposes it appears to be to their mutual benefit if one reads the one in light of the other. While chapter 1 discussed some of the relevant differences between their positions, the continuities in their work are manifest in the present construal of the demarcationist repertoire.

85 In other words: Popper is not Lakatos and though he uses the term 'research programme' he does not seriously consider evolutionary biology holistically in terms of hard core and protective belt. Instead, he takes one by one the explanatory and predictive claims of the theory of evolution by Natural Selection, while on the other hand he conflates the theory of Natural Selection (as an explanatory account *for* the theory of evolution) *with* the theory of evolution. His treatment of explanatory claims on a one by one basis is exemplified by his of the examples that were just cited. His conflated view of the theory of evolution by Natural Selection is manifest in his summary of (modern) Darwinism in his [1976], p. 170:

> (1) The great variety of the forms of life on earth originate from very few forms, perhaps even from a single organism: there is an evolutionary tree, an evolutionary

history.

(2) There is an evolutionary theory which explains this. It consists in the main of the following hypotheses.

(a) Heredity: [...]

(b) Variation: [...]

(c) Natural Selection: there are various mechanisms by which not only the variations but the whole hereditary material is controlled by elimination. [...]

(d) Variability: although *variations* in some sense - the presence of different competitors - are for obvious reasons prior to selection, it may well be the case that *variability* - the scope of variation - is controlled by Natural Selection [...]

Now, compare this summary with the reduction sentence for 'Natural Selection' which - translated into Popperian code - would read:

$$(2a \ \&\ 2b) \Rightarrow [2c \Leftrightarrow (1 \ \&\ 2d)]$$

Elaboration of the 'various mechanisms' mentioned under 2c (as well as elaboration of the principle controlling 2a and 2b) would result from an inquiry into 'Natural Selection' as a placeholder term. This in turn investigation may yield the theory independently testable. As far as the purely internal relations within the reduction sentence are concerned, explanation of 1 and 2d in terms of Natural Selection (2c) does, of course, appear 'merely tautological'.

86 Compare Wassermann [1978] on the testability of Natural Selection in contemporary population genetics.

87 Popper [1978], pp. 345f.

88 For a scathing criticism of Popper's [1976] treatment of Darwin's theory see Ruse [1981]. It appears that Popper's later reflections on the theory of natural selection has not yet encountered similar critical attention.

89 Hooker [1983], p. 85.

90 Carpenter [1983], pp. 113f.

91 Fawcett [1983], pp. 281 and 284.

92 Letter to F.W. Hutton (April 20th, 1861), Darwin [1903], Vol. 1, pp. 183f. (letter # 124). See also the letter to J.D. Hooker (April 23rd, 1861) in Darwin [1887/1893], Vols. 2, p. 362/p. 155.

> He [Hutton] is one of the very few who see that the change of species cannot be directly proved, and that the doctrine must sink or swim according as it groups and explains phenomena.

It is really curious how few judge it in this way, which is clearly the right way.

93 Hull [1983], pp. 27ff. interprets Mill's remark and concludes that Darwin, Fawcett, and Huxley overrated it when they took it for approval of the theory. On this account, however, it appears that Mill said *just what they wanted to hear*: the theory of natural selection is not a proven theory, it is a right step towards the development of a new science.

94 Mill [1874], p. 328. – Note that Mill's 'hypothesis' is Peirce's 'abduction', and that both use 'induction' in the same manner, i.e. as the last stage in the development of a theory.

95 Darwin [1959/1962], XIV, 183/pp. 476f.

96 Here emerges a 'demarcationist' philosophy of science which is informed by the history of science in a way that neither Popper's nor Lakatos's methodologies are. For a sketch of its outline see Shapere [1984], p. xxiii:

> The development of science thus consists in a gradual discovery, sharpening, and organization of relevance-relations, and thus in a gradual separation of the objects of its investigation and what is directly relevant thereto from what is irrelevant to those investigations: a gradual demarcation, that is, of the scientific from the non-scientific. [...] In other words, *this* is what we have come to *call* "scientific". In that development science aims at becoming, as far as possible, autonomous, self-sufficient, in its organization, description, and treatment of its subject-matter

97 Lakatos [1978b], p. 49.

98 Lakatos [1978b], p. 52:

> Which problems scientists working in powerful research programmes rationally choose, is determined by the positive heuristic of the programme rather than by psychologically worrying (or technologically urgent) anomalies. The anomalies are listed and shoved aside in the hope that they will turn, in due course, into corroborations of the programme.

99 Shapin and Barnes [1979], p. 127.

[100]Young [1973], pp. 386f.

[101]Young [1973], p. 394.

[102]Marx [1978], pp. 65f.

[103]Young [1973], p. 385. Young continues on p. 386 by elaborating the methodological foundation of the present remark:

> a radical approach requires that the socio-political basis and its interaction between the putatively autonomous scientific results be explored in depth and detail. He must make this effort in order to understand the role of scientific rationality and its technological expressions (and affiliations) in maintaining the established order of society and in sustaining the false consciousness which prevents men from believing that it can be transformed into a society in which the division of labour need not be hierarchical and exploitative, one in which inegalitarian structures are no longer maintained by being mystified and justified by a spurious foundation in the laws of nature.

Incidentally, Young's political critique of science appears more timely now in respect to Darwin's socio-biologist heirs than to Darwin himself and his theory of natural selection.

[104]Young [1971], p. 500.

[105]For the latter position, see e.g. Pohrt [1972]. In this context, it may be interesting to note that Darwin's theory was appropriated not only by social Darwinists, but also by socialists. While some inferred from the theory of natural selection that government should not interfere with the beneficial economic, cultural, and biological struggle for existence, others held that government *should* interfere to remove the economically and culturally induced inequities in the natural, biological struggle towards improvement of the human race. On the variety of ideological appropriations of Darwin's theories see e.g. Kelly [1981], p. 7.

[106]Shapin and Barnes [1979], p. 128.

[107]Ibid.

[108]Shapin and Barnes [1979], p. 138.

[109]Ibid.

110 Shapin and Barnes [1979], p. 139.

111 Shapin and Barnes [1979], p. 136.

112 Ironically, this is due to the constraints and heuristic power of Shapin and Barnes's own vocabulary of motive. As Mills [1940], p. 912 observes:

> To converted individuals who have become accustomed to the critical terminology of questioning the relative institutional autonomy of science, all others seem self-deceptive.

Of course, this passage is adapted to the present context. Mills *really* talked of individuals "accustomed to the psychoanalytic terminology of motives".

113 Mills [1940].

114 Mills [1940], pp. 909f. - For a critical philosophical discussion of how to assess 'deep interpretations' (as e.g. by Marxism, psychoanalysis etc.), see Danto [1981]: he provides a framework for the argument that deep interpretation in reference to (hidden) reasons is a questionable undertaking, but that its rejection cannot extend to the indispensable interpretation of human action in reference to surface-reasons.

115 Mills [1940], p. 907. In a note on the same page Mills clarifies that

> I am here concerned more with the social function of pronounced motives, than with the sincerity of those pronouncing them.

This clarification establishes the consistency of his view on the efficacy of the vocabularies with the view on the impossibility of explaining social action in terms of what is uttered within any such vocabulary.

116 In this spirit, Karin Knorr-Cetina presents the opportunism of scientists not as *mere* opportunism but as *reasoned* opportunism in the first chapter of her [1981], p. 23:

> In order to realise our interest in the scientists' "cognitive" concerns (rather than their social relations), we must view actual laboratory activities *indiscriminately*. To grasp the meaning of those activities, we must engage ourselves in laboratory reasoning, which reveals the scientist to be a *practical reasoner* who refuses to split into social and technical personalities. What emerges from this reasoning is the *practices* of knowledge

> production, and not some abstract social or cognitive ingredients. The question of how knowlegde is produced and reproduced asks nothing more (and nothing less) than a theory of such practices.

However, her adoption of an ethnographic approach which quite discriminately chooses the laboratory (as opposed to e.g. the scientific journal) as the setting in which science takes place leads to the neglect of this integrative approach when it comes to analyzing the *publicly* available vocabularies of motive. On pp. 22f. for instance, the distinction between social and scientific as employed in the discourse of scientists is recognized "as a resource of strategic interaction". But that it is a resource of strategic interaction (i.e a vocabulary of motive or argumentation repertoire) is then used to challenge the validity of that very distinction. Instead, one should ask how this resource can be utilized by a practical reasoner in order to integrate social *and* institutional demands, to integrate personal goals *and* the constraints that have to be met if there is to be certifiable scientific progress towards objective knowledge. Knorr-Cetina's ethnographic confinement to the setting of the laboratory leads to the following carricature of her work in Rom Harre's preface on p. viii:

> There are many surprises that await us if we enter a laboratory and study a group of scientists in this frame of mind. The idea that the enterprise can be defined in terms of an idealized epistemology, whether that of experimentally based inductions or of the conjectures and empirical refutations of the logicist philosophers of science, is quickly refuted. Logic, it seems, is not among the "idols of the tribe". Where it appears it is an insert in the pursuit of rhetorical advantage in debate.

117 To be sure, this is more easily said than demonstrated. And this is where the philosophical task of certifying the legitimacy of scientific repertoires comes into play. - A few word on this task and its relation to science and scientific change can be found in the conclusion.

118 Mills [1940], p. 913.

119 Merton [1968], pp. 475f. - While Merton points out that the Thomas theorem has a long tradition (he mentions Mandeville, Marx, Hegel, Freud, and Sumner among others who set it forth), the particular challenge posed by the

Thomas theorem to philosophers seems to be this: *not* to see it as a mere re-iteration of notions commonly associated with idealist philosophies of various brands, but instead to understand it as a genuinely sociological theorem on cumulatively effective patterns of social action. As such, it does not at all depend on idealistic (or, for that matter, materialistic) philosophy. It may be useful to remember that Merton's discovery of the 'self-fulfilling prophecy' provides the most conspicuous exemplification of the Thomas theorem.

[120]"The Normative Structure of Science" in Merton [1973], pp. 267-278. It was originally published in 1942 under the title "Science and Technology in a Democratic Order".

[121]Merton [1973], p. 269.

[122]Merton [1973], especially pp. 383-412.

[123]Mulkay [1976], pp. 643f. and 645.

[124]Mulkay [1976], p. 654. It should not surprise that these questions, and particularly the first one, are here labeled 'intriguing' - after all, the present study is trying to lay some groundwork towards answering them.

[125]To be sure, there is a deeper disagreement underlying this passage. It is a disagreement on the use of the words 'norm' and 'ideology'. Mulkay (unlike Merton) seems to expect that norms hold only if they "govern social interaction in [a] straightforward fashion" and if they "are institutionalised within the scientific community in such a way that general conformity is maintained" ([1976], pp. 644 and 654). - Incidentally, Mulkay's tendency (shared with Knorr-Cetina and Shapin and Barnes) to discount the normative vocabulary as 'mere ideology' and 'verbalization' may be the reason for his failure to pursue the research programme which he himself suggested. Gilbert and Mulkay [1984] present a greatly impoverished analysis of scientists' discourse by simplistically relating all utterances to just two repertoires which are vaguely defined as 'empiricist' (formal, internal) and 'contingent' (informal, external). In a perceptive review of their recent work, Shapin [1984] also laments the degeneracy of a research programme which looked so promising in Mulkay's [1976].

[126]That is, their use is only constrained by (L*) (see above, p. 185).

[127]This may sound trivial. It is worth exploring, however, which institutions and professional roles besides science *are* delimited by acceptable argumentation repertoires. Using one of Mills's examples, one may say that diplomats are closely tied to the use of acceptable vocabularies. But in the case of diplomats, those acceptable vocabularies are generated elsewhere (e.g. 'world politics', 'global conflicts', 'presidential politics' etc.) and in a diplomat's experience they belong probably to the most changeable of all variables (and therefore closely observed) - while the institutional role of the diplomat is tied only to certain objectives, commitments of loyalty, interests of state, etc.

[128]This account of the episode is taken from Ellegard [1958], p. 68.

[129]Fawcett [1983], p. 284. - Fawcett's criticism of the Bishop's argument which rested on (nothing but) his own personal authority can be juxtaposed to Fawcett's appeal in the same article to the personal authority of scientists who have a sustained record of adherence to organized scepticism):

> The eminently high authorities who have already welcomed Mr. Darwin's theory as a probablt hypothesis, should induce the general public to welcome it as a legitimate step towards a great scientific discovery (p. 282)

Fawcett's argument is drawn from the professionalist argumentation repertoire: clearly, it cannot (by itself) be used to persuade anybody of the truth of Darwin's theory (and apparently is not meant to do so), but it provides a cogent reason to take it seriously as a scientific theory (implying, as it does, that these authorities have subjected it to demarcationist considerations and judged it to be sufficiently scientific). Accordingly, it would be misleading to think that the norm of organized scepticism should outrule all appeals to authority. On the contrary, it validates some for some contexts, it outrules others in other contexts. (For a differentiated account of authority in science see Polanyi [1964], pp. 203ff.)

[130]As opposed to the tentative definitions of 'scientific activity' in chapter 2, this definimition may hold only for limited periods in the history of science. One should suspect, however, that in its skeletal structure (which is stated here) it holds for science since the latter part of the 17th century.

[131]To John Scott (June 6th, 1863) in Darwin [1903],

Vol. 2, p. 323 (letter # 646).

[132]To W.H. Harvey (August 1860) in Darwin [1903], Vol. 1, p. 160 (letter # 110). - One might look at Darwin's correspondence as a way of making sure that those "others" always knew how to best defend him.

[133]See Hull [1983].

[134]Again, see Hull [1983], and Ellegard [1958].

[135]To Asa Gray (July 22nd, 1860) in Darwin [1887/1893], Vols. 2, p. 327/p. 120.

[136]Hutton [1983], p. 293.

[137]According to Popper [1962], p. 240, it will thus emerge *especially* if it was rejected:

> *De mortuis nil nisi bene*: once a theory is refuted, its empirical character is secure and shines without blemish.

Conclusion: Perspectives II

Aside from its methodological virtues, the theory suggested and expounded here recommends itself as a progressive programme as it explains the five phenomena discussed in the last chapter[1]. Undoubtedly, before the issues of its acceptance as a viable research-programme in the science of science can even arise, a long process of negotiation would have to set in and it should include negotiations between the author and himself. These negotiations will focus mainly on issues and problems which have not so far been addressed at all. And those, in turn, fall into two categories: reflecting on the defects of the programme as stated so far, there are classical philosophical problems which appear to have no clear relation to what has been said here; and reflecting the heuristic promise of the programme, there are a number of clarifications which may emerge from a sustained exploration of this theory of science.

Turning to the defects first, one has to note that the decision-theoretic reduction of 'rationality' does not lead to a sufficiently comprehensive understanding of progress[2] and consensus-formation in science, and indeed, not even of scientific ideas during specific periods in the history of science. While this theory provides a powerful tool for the internal rational reconstruction of decisions that have been made either by individual scientists or the scientific community as a whole, it does not fully explain why a consensus concerning the adoption of 'natural selection' has formed[3]. After all, on the account given so far, one should assume that its opponents made equally productive use to their ends of the available argumentation repertoires. The arguments put forth by Darwin and his compatriots obviously aimed for persuasion. And since in the long run they have actually succeeded in persuading the scientific community, one or both of the following two accounts can be suggested. Either, as Darwin himself predicted, the opponents to 'natural selection' slowly died out and a new generation of more favorably inclined scientists took their place[4]. Or, at some point the opponents of 'natural selection' found themselves in a situation where the price for the continued construction of counter-arguments became higher than letting themselves be persuaded. Both these hypotheses rest on a presupposition which has (tacitly) been made all along and which came to the fore at the very end of the preceding chapter: in principle, human ingenuity and intelligence is powerful enough to indefinitely support any choice among scientific alternatives - even if restricted to acceptable arguments. Now, if one were to adopt the view that only a generational shift is ultimately responsible for the consensus on 'natural selection', one invites a purely externalist interpretaton of consenus-formation: the formation of consensus as well as sustained dissens has to be attributed to the distribution of various goals in the scientific community as the only efficacious variable in scientific negotiation. In contrast, the assumption that consensus is formed when the prize exacted for the construction of counter-arguments becomes too high, invites a mixed external/internal account of consensus-formation. Intuitively, it appears more plausible since it is based on the following two observations. Firstly, human ingenuity may indeed be boundless when it comes to providing an acceptable rationale for whatever choices one makes, but tolerance (one's own as well as that of others) with the contrivance and effort required for sustained dissens is far more limited. And secondly, the connection between some choice among alternatives in science and any given goal is rather tenuous: some religious scientists may initially think that the truth of Darwin's theory somehow contradicts their beliefs concerning the existence of God, and they may therefore tenaciously produce acceptable counter-arguments - but they may find that at some point it becomes preferrable to credibly realign their ideas about God in relation to natural law[5]. Whatever constrains the tolerance of the scientific community and subsequently leads to a realignment in the (overall non-scientific) value-system of those who hitherto stood in the way of consensus, it must be some internal standard governing the give and take of scholarly exchange. For certain situations, this internal standard may be a simple agreed-upon rule which is quite easily discernible, e.g., when all parties have tacitly made the commitment that the outcome of some experiment shall decide over acceptance or rejection of a given hypothesis. In many other situations, however, these internal standards can at best be described in reference to the relative 'beauty' or 'simplicity' of the respective sets of arguments for and against some option[6].

But there is another set of issues associated with consensus-formation and progress in science and it cuts across the external and external/internal proposals which were just presented. It concerns the (certified) argumentation repertoires of science. In these pages, the employment of arguments from a repertoire was equated with the temporary adoption of a methodological posture. However, more general epistemological postures underlie and generate these repertoires. These in turn give content to the notions of 'truth', 'objectivity', and 'rationality'. And particular normative conceptions of 'rationality' in relation to 'truth' and 'objectivity' are particular interpretations of the purely formal decision-theoretic model of rational conduct. If understanding science means understanding the argumentation repertoires available to scientists, much more needs to be learned about the normative features of argumentation repertoires[7]. Now, normative conceptions of 'truth' and 'rationality' change in the course of history and so does the set of available argumentation repertoires. Thus, this deeper issue of scientific rationality corresponds to the historiographical task of determining the scope at any given time of the available argumentation repertoires and the changes in the set of these repertoires. That is, historical inquiry has to determine what at any time is included in the set of argumentation repertoires - which presupposes that there are methods for ascertaining the scope of content of any repertoire at any time. And further, historical inquiry has to unearth the principles and conditions under which the set of argumentation repertoires changes[8]. Answers to these questions would elucidate the character and efficacy of paradigms - and may bring 'historical comprehension' into the reach of the science of science[9].

The programme presented here thus needs to be amended as one moves from the problem of how to rationally account for a given decision to the problem of consensus-formation in science, that is, from the question of how to justify a selection among alternatives in science to the question of why most members of the scientific community end up making the same selection. However: when, under which conditions, and to what end such amendment is actually called for, remains an open question. The reflections of the past four chapters were designed to be indifferent to the answer of this last question.

And yet, even if - for now - one rests content with a decision-theoretically narrowed notion of rationality, fundamental issues clamor for attention. They concern continuities of all scientific work throughout history and independently of paradigm and contemporary varying conceptions of rationality. For instance, it is of foremost interest whether one can discern in the history of science certain algorithms governing the choice of arguments from available repertoires and their use, algorithms that are more specific than the law of tempered opportunism (L*).

But it is time to turn from the more or less pressing gaps or defects to the heuristic merits of the theory. In the course of inquiry, a rather general image of science, scientists, and scientific activity has emerged. If only metaphorically, it can be compared to other such images that are currently engaging the philosophy of science.

Popper views scientists as submitting themselves heroically to the excruciatingly relentless method of trial and error: never claiming victory, they never cease in their critical activity. Popper thus celebrates an aspect of the human intellect that is endemic to good philosophy and good politics, and which is the prerequisite for liberation and change as conceived of by the Enlightenment. Inductivists, on the other hand, are often held to the view that scientists busily and tenaciously collect, sort, patch together and stack building block upon building block slowly upwards into lofty heights. The idea of progress and growth of knowledge is modeled on organic rather than revolutionary processes. A bee-hive comes to mind rather than a Faustian exercise in ultimate futility as one pictures the scientific community through inductivist rather than Popperian spectacles. Lacking the grand gesture towards liberation and enlightenment, these scientists seem a bit more phlegmatic and *gemütlich* or *petite bourgeois*. Yet, here too comes to the fore an important feature of the human intellect, namely, that theories are adaptable to the world, and that patience and endurance (including stubborness based on belief rather than critical attitude) will provide reliable adaptations. Yet another view, however, presents scientists as well-disguised, deceiving and deceptive conspirators working to perpetuate interests and ideologies under the cloak of professed objectivity. Driven by class interests which are not reflected upon within science itself, the scientist satisfies those interests by instinct rather than wit. Scientific method and the relative autonomy of the social institution of science allow scientists to express whatever they please in seemingly objective terms. This view also captures (sad but) important truths about science and the human intellect: namely, the willingness to rest content with localized patches of liberty and ideal discourse. Insisting on the utmost rationality of argumentation within specialized scientific domains, the institution of science readily concedes any decisions *about* science to outside forces which have to withstand no particular scrutiny regarding their rationality. Scientists do not recognize their option to place decisions concerning e.g. the 'finalization' of science within their own domain of rational adjudication. This most cynical and bitter view of science presents crazed geniuses in laboratories within ivory towers insulated from and blind towards a society which masterfully manipulates them. Science, so it seems, owes its relative autonomy only to the expediencies involved in selective appropriation from without of autonomously generated results.

Of course, there is a good deal of overlap between those images of science and the parties who hold them. And certainly, the political affiliations ascribed to them are by no means so very clear-cut. For instance, in historical terms the predicate 'revolutionary' is more accurately applied to some of the early positivists than to mainstream good liberal Karl Popper who shies back from anything smacking of dogma. And a cynical view about the role which scientists play in relation to ideology and society is deeply committed to the value of critical thinking and the confidence that it can and will instigate political change - after all, the complaint of those who hold this view implies that matters do not have to be as they are.

Thus, the image of opportunistic science which emerges from the discussions here, need not contradict any of the extant images of science. Indeed, one might argue that it embodies the more important traits of the other three. Opportunism requires the ability to find means towards an end within an environment that imposes sanctioned restrictions on the available means. As such, the image of opportunistic science celebrates the cunning of reason, *Mutterwitz*, and strategic thinking which simultaneously formulates, questions, and implements the rules of the game: that is, the intellectual capacity to construct working theories of the world upon almost any foundation. Whether opportunism (in that respect quite like guerilla tactics in political 'warfare') is morally reprehensible or not, depends on the chosen means and ends - it has nothing to do with the ability to *be* opportunistic which is certainly admirable. In the case of science and scientists, the end towards which the available means are opportunistically employed may, in effect, be nothing but the truth. And even if not a single scientist was individually interested in the pursuit of truth, the institution of science is presently set up in such a way that no particular goal can be effectively asserted on the institutional scale: the available means (i.e. the argumentative resources) are so constrained that 'truth' as the real or professed goal is asserted (if only serendipitously) while the realization of any other (ideological etc.) goals remains epiphenomenal, subject to selective appropriation by society at large. To explore the transformations of ideology into objective knowledge and the retransformation of objective knowledge into ideologically useful tenets is yet another task emerging from the course of inquiry so far. The 3-phase-model of theory development proposed by the so-called Starnberg group provides an excellent backdrop for such an investigation[10]. It assumes that ideological input is largest before and during the phase of abduction and concept-formation. Such input may subsequently enter into the (actual) adoption of an attitude towards a newly introduced dispositional concept, but it cannot enter into the communal and public process of justifying that choice. Thus, external influence is small in the second phase of theory-development. The third phase begins when scientific work on a given concept or theory is done, at least for the time being. Here, external forces come into play to the extent that scientists cease to be interested in controlling the use of the concept, i.e. as far as the potential use of that concept leaves the realm of autonomous science and becomes subject to selective appropriation by societal forces. In this phase of 'finalization', science sets free into the power-oriented public market-place of ideas an artifact which was at first infested by the ideology of its creators, but then much purified and distilled during tightly constrained scientific adjudication as to become void of all political character. Objective and innocent, it is now an ideal prey for almost any sort of appropriation. This conjectural scheme for the varying interplay between ideology and scientific objectivity, between the public forum and the largely autonomous institution of science will have to be tested and elaborated. And the story of 'natural selection' as it traversed these different stages will have to be told in full.

The image of science that emerges from the preceding chapters pays tribute to the extraordinary achievements of the human intellect, especially in science, without discounting the very real historical pressures which surely left their mark on science as on the products of any other human institution. Within the context of the philosophy of science, Paul Feyerabend seems to have expounded a view most similar to this one[11]. He, too, stresses the continuity between everyday and scientific decision-making: there is no such thing as 'scientific rationality' which could claim a special status of excellence vis-a-vis folk-opportunism. And the characteristically scientific constraints on the available means in the pursuit of ends, and the specifically scientific way of framing alternatives are historically contingent. It seems that Paul Feyerabend would not try to therefore write them off as *merely* contingent. For surely, science may well owe its remarkable achievements to these constraints on argumentative means and theoretical alternatives. If Feyerabend has prescribed a path which this study has followed only to this limited extent, it is for the moral dimension of his thought which poses a challenge to the philosophy of science that the present work is not ready to address: he seems to say that - given our talent for opportunistic reasoning - a dogmatic institutionalization of a certain set of preferred constraints (as, for instance, in science) is unjustified, debilitating, and reactionary. Concerning science he would then say *firstly* that philosophers overestimate the historical efficacy of these constraints, and *secondly* that one would be better off by either relaxing the hold of science on society (i.e. by proliferating institutions that would rival science) or by relaxing the constraints on permissible arguments within science itself. He may be right. In order to decide this, at the very least one would have to investigate the extent to which science progresses in virtue of sanctioned and effective (though mutable and flexible) constraints on the argumentative means and ends which are available to scientists at any given time[12]. But here again, there are two rather different approaches to be taken. That is, one might firstly conjecture that the very same features which set apart the scientific institutional system of knowledge-acquisition from any other such system *also* validate scientific knowledge. However, one might alternatively suggest that the following three issues should carefully be held apart, i.e. the issues (i) of what makes science a unique institution at any given time in history, (ii) of how and why scientific knowledge appears to be outstandingly excellent, (iii) of how science progresses. Philosophers of science have followed these approaches and provided more or less tacit answers to these questions. Their proposals need to be evaluated in the framework of the theory expounded here. But to be sure, much more will have to be done if one wants to fully meet and maybe deflate Feyerabend's challenge.

For now, the institution of science was considered only as a historical artifact. As such, it was presented as an arena in which the cunning of reason becomes most successfully manifest: in the growth of objective knowledge. But whether there is a lesson to be learned from this, and what it tells us, we will have to see.

Notes to Conclusion

[1] Six phenomena, really, if one counts the brief discussion in the first chapter of priority disputes in science (see above, pp. 34f.). And even seven, if one counts (see above, p. 120) the possibility of rational reconstruction of non-referential scientific theories such as phlogistic chemistry.

[2] Here, that term pertains to the overall progress of science throughout history. Of course, there is no difficulty accounting for progress in relation to the aims of research programmes on the basis of the preceding investigation.

[3] The problem of how to account for consensus-formation in science has been posed and dissected most recently and in grand style by Laudan [1984]. However, his presentation confirms the suspicion that at present only tentative approaches towards the solution of that problem are in sight.

[4] Compare Darwin's declaration towards the end of the *Origin* (Darwin [1959/1962/1859/1981], XIV, 191-195/p. 453/pp. 481f./pp. 477f)

> Although I am fully convinced of the truth of the views given in this volume under the form of an abstract, I by no means expect to convince experienced naturalists whose minds are stocked with a multitude of facts all viewed, during a long course of years, from a point of view directly opposite to mine. It is so easy to hide our ignorance under such expressions as the "plan of creation," "unity of design," &c., and to think that we give an explanation when we only restate a fact. Any one whose disposition leads him to attach more weight to unexplained difficulties than to the explanation of a certain number of facts will certainly reject my theory. A few naturalists, endowed with much flexibility of mind, and who have already begun to doubt the immutability of species, may be influenced by this volume; but I look with confidence to the future, to young and rising naturalists, who will be able to view both sides of the question with impartiality. Whoever is led to believe that species are mutable will do good service by conscientiously expressing his conviction; for only thus can the load of prejudice by which

> this subject is overwhelmed be removed.

5 Concerning *inter alia* Darwin's theory Young [1973], p. 350 takes precisely this

> to be the central preoccupation of the participants in the debate: a fundamental reorientation of the conception of the relations between man, God, nature and society. As the work progressed, this reorientation became more closely defined as a change from mechanistic analogies employed within an explicitly theistic natural theology to the use of organic analogies based on secularized, implicit natural theology.

6 A far more general and thorough epistemological theory is needed to fill in the blanks left by the preceding remarks. For the sake of this theory of science the following constraint on any such theory is highly desirable: epistemological theories should construe the notions of 'truth' and 'objectivity' in such a way that any consensus emerging from rational negotiation in science is taken to express objective truth. So-called realist theories do not meet that constraint (unless they employ a very restrictive notion of 'rationality'). And yet, it seems that presently the terms 'objective' and 'true' have been appropriated for the exclusive use of realists to whom the view expounded here must appear hopelessly 'relativistic'. - Compare the end of chapter 4 (pp. 254f.).

7 These normative conceptions of rationality (and the concomitant views on the aims of science) are manifest in the arguments which are drawn from any particular repertoire. However, they tend to evade empirical inquiry since they become discernible only as one investigates the illocutionary force of scientific arguments (see above, note 41 to chapter 1 on pp. 48f.).

8 Here one would expect to find evidence supporting or disconfirming the view that paradigm-change in science is global or revolutionary rather than local or gradual.

9 Compare the quotation from Mittelstrass [1974] in note 10 to chapter 1 on p. 42 above (and note 16 on pp. 43f.). The philosophical and historical investigation of those negotiations which determine the content of scientific argumentation repertoires for any given time may well be conducted within the constructivist research programme envisioned by Mittelstrass. Aside from the previously cited passages see pp. 116f., 144, and especially p. 136:

> a theory of the history of science must be conceived as part of a theory of historical experience. In contrast to other parts of such a theory, it pertains to a kind of knowledge which during all historical eras functions in *praxis*-stabilizing ways - *praxis*-stabilizing not only in respect to methodology but also in respect to teleology. (eine Theorie der Wissenschaftsgeschichte [muss] als ein Teil einer Theorie der historischen Erfahrung begriffen werden. Im Unterschied zu anderen Teilen einer solchen Theorie beträfe sie ein Wissen, das in jeder historischen Phase praxisstabilisierende Funktionen hat, und zwar nicht nur unter methodologischen, sondern auch unter teleologischen Gesichtspunkten.)

Mittelstrass's very abstract considerations bring the history of science back into the fold of the philosophy of history as conceived in the nineteenth-century. Other post-Kuhnian philosophers of science have embarked upon a similar course. Staying clear of epistemological dogmatism on the one hand, (relativistic) rational reconstruction in terms of ever-changing methodological rules on the other, Stegmüller [1979] and Shapere [1984] utilize highly complex and sophisticated theories on the dynamics of science in order to uncover the efficacy of reason in history. While it is easy to agree with Mittelstrass, Stegmüller, and Shapere (especially Shapere), their ideas are most properly construed as regulative ideals for what the philosophy and historiography of science should at some point achieve: but for the immediate purposes of the science of science, the philosophy of history has to take the backseat to an empirical interest in the methodology of historiography.

10 Böhme et.al. [1972], Böhme et.al. [1973], and Schäfer [1978].

11 It does not come as much of a surprise that Feyerabend has indeed used that term as a label for scientists. Compare his [1981], p. 321:

> A scientist is not a high-priest guarding the preservation of basic laws; he is an opportunist who twists the accomplishments of the past to suit once this and then another end - if he pays attention to them at all. (Ein Forscher ist nicht ein Hohepriester, der über die Erhaltung grundlegender Gesetze wacht; er ist ein Opportunist, der Errungenschaften der Vergangenheit bald für diesen, bald für jenen Zweck zurechtbiegt - falls er sie überhaupt beachtet.)

He supports the use of that label in this passage by referring once again to a statement by Albert Einstein which he had already cited in his [1975], p. 18:

> The external conditions which are set for [the scientist] by the facts of experience do not permit him to let himself be too much restricted, in the construction of his conceptual world, by the adherence to an epistemological system. He therefore, must appear to the systematic epistemologist as a type of unscrupulous opportunist....

[12] The qualification "at the very least" should be noted here. For, any such investigation would at best provide the factual basis on which a moral or political decision has to be made. Into that political decision (and further complicating the issue) would have to enter a construal of 'progress' which is not relativized to the goals of research programmes but determined by what is in the interest of humanity.

Bibliography

Technical Notes. Occasionally, a German text was translated into English by A.N. In this case, the original text in German is reproduced in the notes. Also, this bibliography occasionally indicates that there is an English edition of a work which was here cited from the German. And finally, any changes in quoted passages have been marked with square brackets. All italics are by the original authors.

Agassi [1981] Joseph Agassi. Science and Society. Studies in the Sociology of Science. Dordrecht, Boston (Reidel).

Apel [1975] Karl-Otto Apel. Der Denkweg von Charles Sanders Peirce. Eine Einführung in den amerikanischen Pragmatismus. Frankfurt (Suhrkamp).

Armstrong [1969] David M. Armstrong. "Dispositions are Causes." *Analysis*, Vol. 30, pp. 23-26.

Armstrong [1978] David M. Armstrong. "Beliefs as States." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 411-425.

Austin [1967] J.L. Austin. How to Do Things with Words. Cambridge (Harvard University Press).

Barker [1969] A.D. Barker. "An Approach to the Theory of Natural Selection." *Philosophy*, Vol. 44, pp. 271-290.

Barrett et.al. [1981] Paul Barrett, Donald J. Weinshank, Timothy T. Gottleber (eds.). A Concordance to Darwin's Origin of Species, First Edition. Ithaca, London (Cornell University Press).

Beatty [1982] John Beatty. "The Insights and Oversights of Molecular Genetics. The Place of the Evolutionary Perpective." In: P.D. Asquith and T. Nickles (eds.). PSA 1972. East Lansing (Philosophy of Science Association), Vol. 1, pp. 341-355..

Böhme et.al. [1972] Gernot Böhme, Wolfgang van den Daele, Wolfgang Krohn. "Alternativen in der Wissenschaft." *Zeitschrift für Soziologie*, Vol. 1, pp. 302-316.

[English Translation: "Alternatives in Science." *International Journal of Sociology*, Vol. 8, pp. 70-94.]

Böhme et.al. [1973]
Gernot Böhme, Wolfgang van den Daele, Wolfgang Krohn. "Die Finalisierung der Wissenschaft". *Zeitschrift für Soziologie*, Vol. 2, pp. 128-144. [English Translation: "Finalization in Science." *Social Science Information*, Vol. 15, 1976, pp. 307-330.]

Bowler [1983] Peter J. Bowler. The Eclipse of Darwinism. Anti-Darwinian Evolution Theories in the Decades around 1900. Baltimore, London (John Hopkins University Press).

Brandon [1982] Robert Brandon. "The Levels of Selection." In: In: P.D. Asquith and T. Nickles (eds.). PSA 1972. East Lansing (Philosophy of Science Association), Vol. 1, pp. 315-324.

Bridgman [1938] P.W. Bridgman. "Operational Analysis." *Philosophy of Science*, Vol. 5, pp. 114-31.

Campbell [1983] Margaret Campbell. "Adaptation and Fitness." *Studies in History and Philosophy of Science*, Vol. 14, pp. 59-65.

Carnap [1936/37]Rudolf Carnap. "Testability and Meaning." *Philosophy of Science*, Vol. 3, pp. 420-468 and Vol. 4, pp. 1-40.

Carnap [1956] Rudolf Carnap. "The Methodological Character of Theoretical Concepts." In: H. Feigl and M. Scriven (eds.). The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Minneapolis, (University of Minnesota Press), pp. 38-76.

Carnap [1963] Rudolf Carnap. "Replies and Systematic Expositions." In: P.A. Schilpp (ed.). The Philosophy of Rudolf Carnap. La Salle (Open Court), pp. 859-1013.

Carpenter [1983]Willam Benjamin Carpenter. "Darwin on the Origin of Species." In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 88-117.

Cartwright [1983]
Nancy Cartwright. How the Laws of Physics Lie.

Oxford (Oxford University Press).

Coder [1969] David Coder. "Some Misconceptions about Dispositions." *Analysis*, Vol. 29, pp. 200-202.

Cummins [1975] Robert Cummins. "Functional Analysis." *The Journal of Philosophy*, Vol. 72, pp. 741-765.

Danto [1965] Arthur C. Danto. Analytical Philosophy of History. Cambridge (Cambridge University Press).

Danto [1981] Arthur C. Danto. "Deep Interpretation." *The Journal of Philosophy*, Vol. 78, pp. 691-706.

Darwin [1859] Charles Darwin. @UX(The Origin of Species by Means of Natural Selection, or the Prsevertaion of Favoured races in the Struggle for Life.) London (John Murray).

Darwin [1887] Charles Darwin. The Life and Letters of Charles Darwin, including an Autobiographical Chapter. F. Darwin (ed.). 3 vols., London (John Murray).

Darwin [1893] Charles Darwin. The Life and Letters of Charles Darwin. Including an Autobiographical Chapter. F. Darwin (ed.). 2 vols., New York (Appleton).

Darwin [1903] Charles Darwin. More Letters of Charles Darwin. A Record of his Work in a Series of Hitherto Unpublished Letters. F. Darwin (ed.). 2 vols., New York (Appleton). [cited by page and letter number]

Darwin [1958] Charles Darwin. The Autobiography of Charles Darwin. 1809-1882. With original omissions restored, N. Barlow (ed.). New York (Harcourt, Brace).

Darwin [1959] Charles Darwin. The Origin of Species By Charles Darwin. A Variorum Text. M. Peckham (ed.). Philadelphia (University of Pennsylvania Press), 1959. [cited by chapter and sentence numbers]

Darwin [1962] Charles Darwin. The Origin of Species by Means of Natural Selection, or the Prsevertaion of Favoured races in the Struggle for Life. New York, London (Collier). [reprint of 6th edition]

Darwin [1980] Charles Darwin. Metaphysics, Materialism, & the Evolution of Mind. Early Writings of Charles Darwin. Transcribed and annotated by P.H. Barrett. Chicago, London (University of Chicago Press).

Darwin [1980b] Charles Darwin. "Observations on the Parallel Roads of Glen Roy, and of Other Parts of Lochaber in Scotland, with an Attempt to Prove that They Are of Marine Origin." In: P.H. Barrett (ed.). The Collected Papers of Charles Darwin. Chicago, London (University of Chicago Press), Vol. 1, pp. 89-137.

Darwin [1981] Charles Darwin. @UX(The Origin of Species by Means of Natural Selection, or the Prsevertaion of Favoured races in the Struggle for Life.) Harmondworth (Penguin). [reprint of 1st edition]

Davidson [1980] Donald Davidson. Essays on Actions and Events. Oxford (Clarendon Press).

Duhem [1954] Pierre Duhem. The Aim and Structure of Physical Theory. Princeton (Princeton University Press).

Ellegard [1958] Alvar Ellegard. Darwin and the General Reader. The Reception of Darwin's Theory of Evolution in the British Periodical Press, 1859-1872. Göteborg (Acta Universitatis Gothoburgensis. Göteborgs Universitets Arsskrift, Vol. LXIV).

Engelhardt and Zimmermann [1982] Wolf von Engelhardt and Jörg Zimmermann. Theorie der Geowissenschaft. Paderborn, München (Schöningh).

Essler and Trapp [1978] Wilhelm K. Essler und Rainer Trapp. "Some Ways of Operationally Introducing Dispositional Predicates with Regard to Scientific and Ordinary Practice." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 109-134

Fann [1970] K.T. Fann. Peirce's Theory of Abduction. Martinus Nijhoff (The Hague).

Farrall [1975] Lyndsay Farrall. "Controversy and Conflict in Science: A Case Study - The English Biometric School and Mendel's Laws." *Social Studies of Science*, Vol. 5, pp. 269-301.

Fawcett [1983] Henry Fawcett. "A Popular Exposition of Mr. Darwin on the Origin of Species." In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 277-290.

Fetzer [1981] James H. Fetzer. Scientific Knowledge. Causation, Explanation, and Corroboration. Dordrecht, Boston (Reidel).

Feyerabend [1978] Paul Feyerabend. Against Method. Outline of an Anarchistic Theory of Knowledge. London (Verso).

Feyerabend [1981] Paul Feyerabend. "Rückblick." In: H.-P. Duerr (ed.). Versuchungen. Aufsätze zur Philosophie Paul Feyerabends. Frankfurt (Suhrkamp), Vol. 2, pp. 320-372.

Fisher [1958] Ronald A. Fisher. The Genetical Theory of Natural Selection. 2nd revised edition. New York (Dover).

Fisk [1978] Milton Fisk. "Capacities and Natures." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 189-210.

Gale [1982] Barry G. Gale. Evolution without Evidence. Charles Darwin and *The Origin of Species*. Albuquerque (University New Mexico Press).

Garber [1983] Daniel Garber. "Old Evidence and Logical Omniscience in Bayesian Confirmation Theory." In: J. Earman (ed.). Testing Scientific Theories. Minneapolis (University of Minnesota Press), pp. 99-132.

Gethmann [1980] Carl Friedrich Gethmann. "Die Logik der Wissenschaftstheorie." In: C.F. Gethmann (ed.). Theorie des wissenschaftlichen Argumentierens. Frankfurt (Suhrkamp), pp. 15-42.

Ghiselin [1984] Michael T. Ghiselin. The Triumph of the Darwinian Method. Chicago, London (University of Chicago Press).

Gilbert and Mulkay [1984] G. Nigel Gilbert and Michael Mulkay. Opening Pandora's Box. A Sociological Analysis of Scientists's Discourse. Cambridge, London (Cambridge University Press).

Glymour [1981] Clark Glymour. Theory and Evidence. Princeton (Princeton University Press).

Goodman [1973] Nelson Goodman. Fact, Fiction, and Forecast. Third Edition. Indianapolis, New York (Bobbs-Merrill).

Gray [1963] Asa Gray. Darwiniana. Essays and Reviews Pertaining to Darwinism. Cambridge (Harvard University Press).

Gruber [1981] Howard Gruber. Darwin on Man. A Psychological Study of Scientific Creativity. Chicago, London (University of Chicago Press).

Habermas [1981] Jürgen Habermas. Theorie des kommunikativen Handelns. Vol. 1, Frankfurt (Suhrkamp).

Hall [1971] Richard Hall. "Can We Use the History of Science to Decide between Competing Methodologies?" In: R.C. Buck and R.S. Cohen (eds.). PSA 1970. In Memory of Rudolf Carnap. Dordrecht, Boston (Reidel), pp. 151-159.

Hanson [1961] Norwood Russell Hanson. Patterns of Discovery. An Inquiry into the Conceptual Foundations of Science. Cambridge (University of Cambrige Press).

Haughton [1983] Samuel Haughton. "Biogenesis." In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 217-227.

Hempel [1963] Carl G. Hempel. "Implications of Carnap's Work for the Philosophy of Science." In: P.A. Schilpp (ed.). The Philosophy of Rudolf Carnap. La Salle (Open Court), pp. 685-709.

Hempel [1965] Carl G. Hempel. Aspects of Scientific Explanation. And other Essays in the Philosophy of Science. New York (The Free Press).

Hempel and Oppenheim [1965]
Carl G. Hempel and Paul Oppenheim. "Studies in the Logic of Explanation." In: C.G. Hempel. Aspects of Scientific Explanation. And other Essays in the Philosophy of Science. New York (The Free Press), pp. 245-290.

Himmelfarb [1959]
Gertrude Himmelfarb. Darwin and the Darwinian Revolution. London (Chatto & Windus).

Hooker [1983] Joseph Dalton Hooker. "Review of the *Origin of Species*". In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 81-85.

Hopkins [1983] William Hopkins. "Physical Theories of the Phenomena of Life." In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 229-275.

Hull [1983] David L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press).

Hutton [1983] Frederick Wollaston Hutton. "Review of the Origin of Species." In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 292-300.

Huxley [1906] Thomas Henry Huxley. Man's Place in Nature and Other Essays. London (J.M. Dent).

Huxley [1942] Julian Huxley. Evolution. The Modern Synthesis. London (Allen & Unwin).

Kant [1911] Immanuel Kant. Kant's Critique of Aesthetic Judgement. J.C. Meredith (ed.). Oxford (Clarendon Press).

Kelly [1981] Alfred Kelly. The Descent of Darwin: The Popularization of Darwinism in Germany, 1860-1914. Chapel Hill (University of North Carolina Press).

Knorr-Cetina [1981]
Karin Knorr-Cetina. The Manufacture of Knowledge. An Essay on the Constructivist and Contextual Nature of Science. Oxford, New York (Pergamon Press).

Kuhn [1970] Thomas S. Kuhn. The Structure of Scientific Revolutions. Second, enlarged Edition. Chicago (University of Chicago Press).

Kuhn [1971] Thomas S. Kuhn. "Notes on Lakatos." In: R.C. Buck and R.S. Cohen (eds.). PSA 1970. In Memory of Rudolf Carnap. Dordrecht, Boston (Reidel), pp. 137-146.

Kuhn [1972] Thomas S. Kuhn. "Reflections on my Critics." In: I. Lakatos and A. Musgrave (eds.). Criticism and the Growth of Knowledge. Cambridge (Cambridge University Press), pp. 231-278.

Kuhn [1983] Thomas S. Kuhn. "Rationality and Theory Choice." *The Journal of Philosophy*, Vol. 80, pp. 563-570.

Küng [1980] Hans Küng. Does God Exist? An Answer for Today. Garden City, New York (Doubleday).

Kyburg [1983] Henry Kyburg. "The Deductive Model: Does it have Instances?" In: J. Earman (ed.). Testing Scientific Theories. Minneapolis (University of Minnesota Press), pp. 299-312.

Lakatos [1978] Imre Lakatos. "History of Science and its Rational Reconstructions." In: I. Lakatos. The Methodology of Scientific Research Programmes. Philosophical Papers Volume 1. Ed. by J. Worrall and G. Currie. Cambridge (Cambridge University Press), pp. 102-138.

Lakatos [1978b] Imre Lakatos. "Falsification and the Methodology of Scientific Research Programmes." In: I. Lakatos. The Methodology of Scientific Research Programmes. Philosophical Papers Volume 1. Ed. by J. Worrall and G. Currie. Cambridge (Cambridge University Press), pp. 8-101.

Lakatos [1978c] Imre Lakatos. Mathematics, Science, and Epistemology. Philosophical Papers Volume 2. Ed. by J. Worrall and G. Currie. Cambridge (Cambridge University Press).

Laudan [1984] Larry Laudan. Science and Values. The Aims of Science and Their Role in Scientific Debate. Berkeley, Los Angeles (University of California Press).

Lee [1969] K.K. Lee. "Popper's Falsifiability and Darwin's Natural Selection." *Philosophy*, Vol. 44, pp. 291-302.

Levi [1967] Isaac Levi. Gambling with Truth. An Essay on Induction and the Aims of Science. Cambridge, London (The MIT Press).

Levi [1969] Isaac Levi. "Are Statistical Hypotheses Covering Laws?" *Synthese*, Vol. 20, pp. 297-307.

Levi [1976] Isaac Levi. "A Paradox for the Birds." In: R.S. Cohen, P.K. Feyerabend and M.W. Wartofsky (eds.). Essays in Memory of Imre Lakatos. Boston, Dordrecht (Reidel), pp. 371-378.

Levi [1978] Isaac Levi. "Subjunctives, Dispositions, and Chances." In: R. Tuomela (ed.). Dispositions.

Dordrecht, Boston (Reidel), pp. 303-335.

Levi [1979] Isaac Levi. "Abduction and Demands for Information." In: I. Niiniluoto and R. Tuomela (eds.). Logic and Epistemology of Scientific Change. *Acta Philosophica Fennica*, Vol. 30 (North-Holland), pp. 405-429.

Levi [1980] Isaac Levi. The Enterprise of Knowledge. An Essay on Knowledge, Credal Probability, and Chance. Cambridge, London (The MIT Press).

Levi [1983] Isaac Levi. "Truth, Fallibility and the Growth of Knowledge." In: R.S. Cohen and M.W. Wartofsky (eds.). Language, Logic, and Method. Boston, Dordrecht (Reidel), pp. 153-174.

Levi and Morgenbesser [1978]
Isaac Levi and Sidney Morgenbesser. "Belief and Disposition." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 389-410.

Lewontin et. al. [1984]
Richard Lewontin, Steven Rose, Leon J. Kamin. Not in Our Genes. Biology, Ideology, and Human Nature. New York (Pantheon).

Lichtenberg [1975]
Georg Christoph Lichtenberg. Schriften und Briefe. Zweiter Band. Sudelbücher II, Materialhefte, Tagebücher. Ed. by W. Promies. München (Carl Hanser).

Lloyd [1983] Elisabeth A. Lloyd. "The Nature of Darwin's Support for the Theory of Natural Selection." *Philosophy of Science*, Vol. 50, pp. 112-129.

MacIntyre [1962] Alasdair MacIntyre. "A Mistake about Causality in Social Science." In: P. Laslett and W.G. Runciman (eds.). Philosophy, Politics and Society. Second Series. Oxford (Blackwell), pp. 48-70.

MacKenzie and Barnes [1979]
Donald MacKenzie and Barry Barnes. "Scientific Judgment: The Biometry-Mendelism Controversy." In: B. Barnes and S. Shapin. Natural Order. Historical Studies of Scientific Culture. Beverly Hills, London (SAGE), pp. 191-210.

Manier [1978] Edward Manier. The Young Darwin and His Cultural Circle. Dordrect, Boston (Reidel).

Manser [1965] A.R. Manser. "The Concept of Evolution."

Philosophy, Vol. 40, pp. 18-34.

Marx [1978] Karl Marx and Friedrich Engels. The German Ideology. Part One. C.J. Arthur (ed.). New York (International).

Mayr [1976] Ernst Mayr. Evolution and the Diversity of Life. Selected Essays. Cambridge, London (Harvard Univesity Press).

Mellor [1978] D. Hugh Mellor. "In Defense of Dispositions." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 55-76.

Merton [1968] Robert K. Merton. Social Theory and Social Structure. Enlarged Edition. New York, London (Free Press).

Merton [1973] Robert K. Merton. The Sociology of Science. Theoretical and Empirical Investigations. Ed. by N. Storer. Chicage (University of Chicago Press).

Merton [1981] Robert K. Merton. "Foreword. Remarks on Theoretical Pluralism." In: P.M. Blau and R.K. Merton (eds.). Continuities in Structural Inquiry. London, Beverly Hills (SAGE Publications), pp. i-vii.

Merton and Zuckerman [1973] Robert K. Merton and Harriet Zuckerman. "Institutionalized Patterns of Evaluation in Science." In: Robert K. Merton. The Sociology of Science. Theoretical and Empirical Investigations. Ed. by N. Storer. Chicage (University of Chicago Press), pp. 460-496.

Michod [1981] Richard E. Michod. "Positive Heuristics in Evolutionary Biology." *British Journal for the Philosophy of Science*, Vol. 32, pp. 1-36.

Mill [1874] John Stuart Mill. A System of Logic, etc. 8th ed. London (Longmans, Green).

Mills [1940] C. Wright Mills. "Situated Actions and Vocabularies of Motive." *American Sociological Review*, Vol. 5, pp. 904-913.

Mittelstrass [1974] Jürgen Mittelstrass. Die Möglichkeit von Wissenschaft. Frankfurt (Suhrkamp).

Moliere [1950] Moliere. The Would-Be Invalid. Transl. by Morris Bishop. New York (Appleton-Century).

Mulkay [1976] Michael J. Mulkay. "Norms and Ideology in Science." *Social Science Information*, Vol. 15, pp. 637-656.

Naess [1972] Arne Naess. The Pluralist and Possibilist Aspect of the Scientific Enterprise. Oslo (Universitetsforlaget), London (George Allen & Unwin).

Nagel [1979] Ernest Nagel. The Structure of Science. Problems in the Logic of Scientific Explanation. Indianapolis, Cambridge (Hackett).

Nagel [1979b] Ernest Nagel. Teleology Revisited and Other Essays in the Philosophy and History of Science. New York (Columbia University Press.)

Newton [1958] Isaac Newton. Papers & Letters on Natural Philosophy and related documents. Ed. by I.B. Cohen [and R. Schofield]. Cambridge (Harvard University Press).

Nordmann [1986] Alfred Nordmann. "Comparing Incommensurable Theories. A Textbook-Account from 1794." *Studies in History and Philosophy of Science* [forthcoming].

Owen [1868] Richard Owen. The Anatomy of Vertebrates. Vol. 3, London (Longmans, Green).

Owen [1983] Richard Owen. "Darwin on the Origin of Species". In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 175-213.

Pap [1963] Arthur Pap. "Reduction Sentences and Disposition Predicates." In: P.A. Schilpp (ed.). The Philosophy of Rudolf Carnap. La Salle (Open Court), pp. 559-597.

Peirce [1958-60] Charles Sanders Peirce. Collected Papers. Volume V: Pragmatism and Pragmaticism. Volume VI: Scientific Metaphysics. Volume VII: Science and Philosophy. Ed. by C. Hartshorne, P. Weiss, and A. Burks. Cambridge (Harvard University Press) [cited by volume and paragraph numbers].

Pictet [1983] Francois Jules Pictet. "On the Origin of Species by Charles Darwin". In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London

(University of Chicago Press), pp. 142-152.

Pohrt [1972] Wolfgang Pohrt. "Skizze zur Entwicklung des Verhältnisses von Wissenschaft und Gesellschaft." In: Wolfgang Pohrt (ed.). Wissenschaftspolitik - von wem, für wen, wie? Prioritäten in der Forschungsplanung. München (Hanser), pp. 45-76.

Polanyi [1964] Michael Polanyi. Personal Knowledge. Towards a Post-Critical Philosophy. New York, Evanston (Harper & Row).

Popper [1962] Karl Popper. Conjectures and Refutations. New York, London (Basic Books).

Popper [1963] Karl Popper. "The Demarcation of Science and Metaphysics." In: P.A. Schilpp (ed.). The Philosophy of Rudolf Carnap. La Salle (Open Court), pp. 183-226.

Popper [1966] Karl Popper. The Open Society and Its Enemies. 2 vols, Princeton (Princeton University Press).

Popper [1968] Karl Popper. The Logic of Scientific Discovery. New York, Hagerstown (Harper & Row).

Popper [1972] Karl Popper. "Normal Science and its Dangers." In: I. Lakatos and A. Musgrave (eds.). Criticism and the Growth of Knowledge. Cambridge (Cambridge University Press), pp. 51-58.

Popper [1975] Karl Popper. Objective Knowledge. An Evolutionary Approach. Oxford (Clarendon Press).

Popper [1976] Karl Popper. Unended Quest. An Intellectual Autobiography. (Fontana/Collins).

Popper [1978] Karl Popper. "The Propensity Interpretation of Probability." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 247-265.

Popper [1978] Karl Popper. "Natural Selection and the Emergence of Mind." *Dialectica*, Vol. 32, pp. 339-355.

Prausnitz [1982] John M. Prausnitz. "Solutions and Solubilities." *The New Encyclopaedia Britannica. Macropaedia*, 15th edition, Vol. 16, pp. 1047-1057.

Provine [1971] William B. Provine (ed.). The Origins of Theoretical Population Genetics. Chicago, London (University of Chicago Press).

Punnett [1911] R.C. Punnett. Mendelism. New York (Macmillan).

Quine [1978] Willard van Orman Quine. "Disposition." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 155-161.

Radnitzky [1971] Gerard Radnitzky. "Theorienpluralismus - Theorienmonismus. Einer der Faktoren, die den Forschungsprozess beeinflussen und die selbst von Weltbildannahmen abhängig sind." In: A. Diemer (ed.). Der Methoden- und Theorienplurlismus in den Wissenschaften. Meisenheim (Anton Hain), pp. 135-184.

Radnitzky [1979] Gerard Radnitzky. "Justfying a Theory Versus Giving Good Reasons for Preferring a Theory. On the Big Divide in the Philosophy of Science." In: G. Radnitzky and G. Anderson (eds.). The Structure and Development of Science. Dordrecht, Boston (Reidel), pp. 213-256.

Reichenbach [1961]
Hans Reichenbach. Experience and Prediction. An Analysis of the Foundations of the Structure of Knowledge. Chicago, London (University of Chicago Press).

Richards [1983] Robert J. Richards. "Why Darwin Delayed, Or Interesting Problems and Models in the History of Science." *Journal for the History of the Behavioral Sciences*, Vol. 19, pp. 45-53.

Ringen [1982] Jon D. Ringen. "The Explanatory Import of Dispositions. A Defense of Scientific Realism." In: P.D. Asquith and T. Nickles (eds.). PSA 1982. Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association. East Lansing (Philosophy of Science Association), Vol. 1, pp. 122-133.

Roll-Hansen [1980]
Nils Roll-Hansen. "The controversy between biometricians and Mendelians: A test case for the sociology of scientific knowlegde." *Social Science Information*, Vol. 19, pp. 501-517.

Romanes [1896] George John Romanes. Darwin, and after Darwin. Vol. I: The Darwinian Theory. Chicago (Open Court).

Rudwick [1974] Martin Rudwick. "Darwin and Glen Roy: A 'Great Failure' in Scientific Method?" *Studies in History and Philosophy of Science*, Vol. 5, pp.

99-178.

Ruse [1971] Michael Ruse. "Natural Selection in *The Origin of Species.*" *History and Philosophy of Science*, Vol. 1, pp. 311-351.

Ruse [1981] Michael Ruse. "Karl Popper and Evolutionary Biology." In: Michael Ruse. Is Science Sexist? Dordrecht, Boston (Reidel), pp. 65-84.

Ruse [1979] Michael Ruse. The Darwinian Revolution. Science Red in Tooth and Claw. Chicago, London (University of Chicago Press).

Ryle [1949] Gilbert Ryle. The Concept of Mind. New York, San Francisco (Barnes & Nobles).

Salmon [1968] Wesley Salmon. "The Justification of Inductive Rules of Inference." In: I. Lakatos (ed.). The Problem of Inductive Logic. Vol 2, Amsterdam (North-Holland), pp. 24-43.

Schäfer [1974] Lothar Schäfer. Erfahrung und Konvention. Zum Theoriebegriff der empirischen Wissenschaften. Stuttgart, Bad Cannstatt (problemata).

Schäfer [1978] Wolf Schäfer. "Normative Finalisierung. Eine Perspketive." In; G. Böhme, W. van den Daele, et.al. (eds.). Starnberger Studien I: Die gesellschaftliche Orientierung des wissenschaftlichen Fortschritts. Frankfurt (Suhrkamp).

Scheffler [1981] Israel Scheffler. The Anatomy of Inquiry. Philosophical Studies in the Theory of Science. Indianapolis, Cambridge (Hackett).

Schnädelbach [1971]
Herbert Schnädelbach. "Dispositionsbegriffe der Erkenntnistheorie. Zum Problem ihrer Sinnbedingungen." *Zeitschrift für allgemeine Wissenschaftstheorie*, Vol. 2, pp. 89-100.

Shapere [1984] Dudley Shapere. Reason and the Search for Knowledge. Investigations in the Philosophy of Science. Dordrecht, Boston (Reidel).

Shapin [1984] Steven Shapin. "Talking History. Reflections on Discourse Analysis." *Isis*, Vol. 75, pp. 125-128.

Shapin and Barnes [1979]
Steven Shapin and Barry Barnes. "Darwin and Social Darwinism: Purity and History." In: B.

Barnes and S. Shapin (eds.). Natural Order. Historical Studies of Scientific Culture. Beverly Hills, London (SAGE), pp. 125-142.

Squires (1968) Roger Squires. "Are Dispositions Causes?" *Analysis*, Vol. 29, pp. 45-47.

Stegmüller [1975] Wolfgang Stegmüller. Das Problem der Induktion: Humes Herausforderung und moderne Antworten. Darmstadt (Wissenschaftliche Buchgesellschaft).

Stegmüller [1979] Wolfgang Stegmüller. "A Combined Approach to the Dynamics of Theories. How to improve Historical Interpretations of Theory Change by applying Set Theoreticl Structures." In: G. Radnitzky and G. Anderson (eds.). The Structure and Development of Science. Dordrecht, Boston (Reidel), pp. 151-186.

Stevenson (1969) Leslie Stevenson. "Are Dispositions Causes?" *Analysis*), Vol. 29, pp. 197-199.

Thompson [1983] Paul Thompson. "The Structure of Evolutionary Theory: A Semantic Approach." *Studies in History and Philosophy of Science*, Vol. 14, pp. 215-229.

Tuomela [1978] Raimo Tuomela. "Dispositions, Realism, and Explanation." In: R. Tuomela (ed.). Dispositions. Dordrecht, Boston (Reidel), pp. 427-448.

Urbach [1978] Peter Urbach. "The Objective Promise of a Research Programme." In: G. Radnitzky and G. Andersson (eds.). Progress and Rationality in Science. Dordrecht, Boston (Reidel), 1978, pp. 99-113.

Van Fraassen [1980] Bastian Van Fraassen. The Scientific Image. Oxford (Clarendon Press).

Wassermann [1978] Gerhard D. Wassermann. "Testability of the Role of Natural Selection within Theories of Population Genetics and Evolution." *British Journal for the Philosophy of Science*, Vol. 29, pp. 223-242.

Wollaston [1983] Thomas Vernon Wollaston. "Review of the *Origin of Species*." In: D. L. Hull. Darwin and His Critics. The Reception of Darwin's Theory by the Scientific Community. Chicago, London (University of Chicago Press), pp. 127-140.

Young [1971] Robert Young. "Darwin's Metaphor: Does Nature Select?" *The Monist*, Vol. 55, pp. 442-503.

Young [1973] Robert Young. "The Historiographic and Ideological Contexts of the Nineteenth-Century Debate on Man's Place in Nature." In: M. Teich and R. Young (eds.). Changing Perspectives in the History of Science. Essays in Honor of Joseph Needham. Dordrecht, Boston (Reidel), pp. 344-438.

Zilsel [1941] Edgar Zilsel. "Physics and the Problem of Historico-sociological Laws." *Philosophy of Science*, Vol. 8, pp. 567-579.

Lebenslauf

Ich kam am 12. November 1956 in Bad Harzburg als Sohn von Otto Nordmann und Katja Nordmann-Mörike zur Welt. Nach meinem Abitur im Jahr 1976 begann ich mein Studium am Leibniz Kolleg und am philosophischen Seminar der Universität Tübingen. Im Herbst 1977 setzte ich in Hamburg das Studium von Philosphie, Germanistik (Neuere Deutsche Literatur) und Geschichte der Naturwissenschaften fort. 1981 erhielt ich den Magisterabschluss für eine Arbeit über "Wissenschaftsbegriff und aphoristische Darstellung bei Georg Christoph Lichtenberg". Nachdem ich bereits das Jahr 1973/74 als Austauschschüler in den USA verbracht hatte, ging ich im Herbst 1981 wiederum für längere Zeit nach Amerika, wo ich als "Visiting Scholar" an der Columbia Universität insbesondere wissenschaftssoziologische Studien unter Robert K. Merton betrieb.

Erklärung

Hiermit erkläre ich, dass ich vorstehende Arbeit selbständig und ohne fremde Hilfe verfasst sowie andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht habe.

Hamburg, den 24. 9. 1985

Alfred Nordmann

(Alfred Nordmann)