THE AESTHETIC EVALUATION OF SCIENTIFIC THEORIES
A dissertation submitted for the degree of Ph.D. of the University of Cambridge

James William McAllister
DARWIN COLLEGE

February 1989
THE AESTHETIC EVALUATION
OF SCIENTIFIC THEORIES

James W. McAllister
DARWIN COLLEGE, CAMBRIDGE

This dissertation studies the incidence of aesthetic criteria of theory-evaluation and theory-choice on the development of science. It suggests that scientific communities customarily appeal to two sets of criteria of theory-evaluation in choosing between competing theories. Criteria of the first set are logico-empirical in nature and are indicative of the observational success of the theories to which they are applied: such criteria refer to such features of theories as their logical consistency and predictive accuracy. Criteria of the second set, which have hitherto received far less attention in philosophy of science, are aesthetic in nature and not indicative of theories' empirical success; they refer to such features of theories as the form of their simplicity or of their symmetry.

The dissertation suggests that the mode of perception under which criteria of the latter set are applied to scientific theories is a mode of disinterested perception, which has no regard for the utilitarian virtues of theories. The conception of the mode of aesthetic perception as disinterested is retraced to its eighteenth-century formulations in the works of Shaftesbury and Hutcheson as well as to some of its twentieth-century proponents. A model of the evolution of aesthetic canons of theory-evaluation is proposed, which suggests that the aesthetic evaluative criteria applied by scientists to their theories are updated to reflect the aesthetic features of past theories which have demonstrated striking empirical success.

These elements of scientific methodology are lastly used to construct a new interpretation of the notion of scientific revolution, as a discontinuous exchange of one aesthetic canon for another. This construal allows the belief in the occurrence of revolutionary discontinuities in scientific methodology to be meshed with a realist and rationalist view of the history of science.

The dissertation contains studies of four episodes in history of science in which the operation of aesthetic criteria of theory-assessment may be discerned. These episodes consist of the formulation and early reception of the astronomical theories of Copernicus and Kepler, of the theory of special relativity, and of theoretical elements of quantum mechanics.

February 1989
CONTENTS

PREFACE iii

Chapter One
THEMES AND PROCEDURES 1
1. Two problems for the rationalist image 1
2. Mutual evaluations of science and philosophy 4
3. The value of scientists’ testimony 7

Chapter Two
THE STIPULATION OF METHODOLOGICAL NORMS 12
1. War games and the philosophy of science 12
2. The constitution of scientific success 17
3. The logico-empirical rule-book 19
4. The recourse to extra-empirical criteria 23

Chapter Three
BEAUTY DISCERNED IN SCIENCE 27
1. Judgements of utility and of aesthetic value 27
2. Hutcheson’s discussion of intellectual beauty 30
3. Some twentieth-century invocations of beauty in science 37
4. The literary reception of scientific theories 39
5. The perception of beauty in abstract objects 43

Chapter Four
THE AESTHETIC STANCE 44
1. Two views of aesthetics 44
2. The study of disinterested attention 48
3. The rise of the discipline of aesthetics 52

Chapter Five
THEORY-EVALUATION ON AESTHETIC GROUNDS 60
1. A conceptual vacancy is announced and filled 60
2. Positivist denials of aesthetic theory-evaluation 64
3. Transcending the positivist demarcation 68
4. The disjunction of aesthetic and epistemic values 71
5. Aesthetic criteria and the underdetermination of theory-choice 77
6. The place of aesthetics in Einstein’s view of theory-assessment 79
7. Alternative justificatory roles of aesthetic features 83
# Contents

**Chapter Six**

Aesthetic Features to Have Won Commendation  
1. The specificity of descriptions of aesthetic canons  86  
2. Aims of the survey  88  
3. Simplicity  89  
4. Symmetry  94  
5. Analogical tractability  97  
6. Consonance with metaphysical presuppositions  102

**Chapter Seven**

The Inductive Construction of Aesthetic Preference  
1. The determinacy of theory-choice  104  
2. The origin and mutability of methodological precepts  106  
3. The aesthetic metainduction  110  
4. Aspects of the evolution of aesthetic canons  115  
5. The psychological motivation of the metainduction  121

**Chapter Eight**

Revolution as Aesthetic Rupture  
1. Gradualism and catastrophism in the philosophy of science  124  
2. Aesthetic canons in normal science  128  
3. The abandonment of aesthetic commitments  134  
4. The aesthetic interpretation of revolutions  138  
5. Quarrels with the Kuhnian view of aesthetic factors  141  
6. Boltzmann’s model of scientific revolutions  145  
7. The analogy with revolutions in art  148  
8. The analogy with political revolutions  150

**Chapter Nine**

Implications for Historiography  
1. The possibility of writing history  152  
2. The recognition of scientific revolutions  154

**Chapter Ten**

Applications of the Model to History of Science  
1. The test of the model  159  
2. Aspects of continuity: Copernicus’s contribution to astronomy  162  
3. Aspects of revolution: the Keplerian ellipses  175  
4. Aspects of continuity: the theory of special relativity  182  
5. Aspects of revolution: quantum mechanics  187

**Chapter Eleven**

The Definition of Scientific Rationality  
1. Review of the argument  192  
2. The scientist’s look  194  
3. The adherence to rationalist precepts  195  
4. The memory of science  198

Bibliography  201
PREFACE

This dissertation is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Cambridge. No part of this dissertation has been submitted for any qualification of any other university. It is in its entirety my own work and includes nothing which is the outcome of work done in collaboration with others. It does not exceed 80,000 words in length.

For invaluable help and advice which they have extended to me, I wish to thank Dr Nicholas Jardine and Prof. Mary B. Hesse of the Department of History and Philosophy of Science, University of Cambridge, who have at different times of the period of my registration supervised my research.

For financial support in aid of this work I thank the British Academy, Jesus College and Darwin College, Cambridge, and the electors to the Gerstenberg Studentship, University of Cambridge.

Earlier versions of some of the arguments pursued here were presented in March 1987 at a seminar at the Department of History of the Université de Montréal and in October 1988 at a seminar at the Department of History and Philosophy of Science at King's College, London. I am grateful to the participants at both seminars for their attention and comments. A brief published version of some of the conclusions reached here will appear under the title 'Truth and Beauty in Scientific Reason' in *Synthese*, 1989.

James W. McAllister
Darwin College, Cambridge
Chapter One
THEMES AND PROCEDURES

Beauty. Publish this surely, beautiful!

Robert Millikan

1. Two problems for the rationalist image

An exhaustive description of scientific methodology which renders full justice to its historical, disciplinary and sheer idiosyncratic variability is a fond philosophical desideratum which it has so far proved impossible to construct. While awaiting its completion, a partial understanding of scientific method is attained by appeal to a set of partial models each of which accounts reasonably well for some facet or other of the explanandum.

The models which compose this set fall into a hierarchy of generality. The models at the top of the structure attempt to account for the most general components of scientific method or for the broadest sweep of the history of science, but are insufficiently detailed to provide explanations of individual elements of each. Models of lower generality account for single methodological precepts or isolated episodes in the historical development of science, but are of little aid to an understanding of the large-scale properties of the enterprise.

The greatest explanatory satisfaction is obtained if the models which compose this set are mutually consistent, so that a glance down the hierarchical structure permits the observer to retain conceptual continuity while homing in from the widest view to the smallest detail of the object of study.

One of the most influential of the models at the top level of this hierarchy has been the rationalist image of science. This model supposes that there exists a set of methodological precepts – the norms of rationality – which admits of some

---

1 Cited in Holton (1978), p. 67; emphasis in the original. This remark, inscribed in Millikan’s laboratory notebook, is his reaction at some of the data collected in March 1912 in the oil-drop experiments to determine the electronic charge. The page of Millikan’s notebook on which the remark appears is reproduced ibid., p. 64; for further information on the experiments and comments on the methodological implications of Millikan’s remark see ibid., pp. 25-83.
impartial and extra-historical justification, against which the actual methodology followed by scientists in history can be compared, and from which this actual methodology did not too persistently or too greatly depart. Many philosophers of science believe that this is a valuable top-level model of scientific method because, among other reasons, there exists a large number of lower-level models which are consistent with it and which provide explanations of reasonably good quality for individual facets of scientific method or particular episodes of the history of science.

In recent years however considerable support has been gained by two explanatory models at a lower level of the hierarchy which seem to be incompatible with the rationalist image of science. This development constitutes a challenge to the rationalist image, since it may presage the day in which another top-level model of scientific method is preferred to the rationalist image on account of its greater consonance with popular lower-level models.

Of the two models now in apparent conflict with the rationalist image, the first – which has been elaborated in considerable detail and which poses the greater challenge – claims that the past development of science has been fractured by revolutions into distinct epochs adhering to norms of theory-formulation and theory-evaluation peculiar to each and not altogether shared by adjacent periods. The second model, of which only elements have hitherto been delineated and which thus mounts the less well-studied threat to the rationalist image, claims that in the history of science many important acts of theory-evaluation and theory-choice have been decided partly on grounds which related not to the observational success of theories but rather to their aesthetic features. On this claim, the development of science over centuries – which is itself no more than the effect of a sequence of acts of theory-choice – is shaped partly by aesthetic considerations and not entirely by calculations of empirical success.

Each of these lower-level models appears to undermine the viability of the rationalist image as the top-level model for scientific methodology, by suggesting that important components of either a systematic or a historicist description of that methodology would conflict with presuppositions held by the rationalist image. The first of the lower-level models described above suggests that over the course of past science theory-formulation and evaluation have been justified by
CHAPTER ONE: THEMES AND PROCEDURES

appeal to an indefinitely large number of different canons of rationality, and furthermore that there exist no grounds upon which to confer privilege to any one of these canons as a basis for a reconstruction of the developments in science in periods other than the one in which appeal to that canon was customarily made. The second model suggests that choices among theories have furthermore commonly been influenced by scientists' preferences in the matter of the aesthetic features of the theories with which they were presented. If these predispositions of scientists are as idiosyncratic and irreducible to rational deliberation as they are generally and tacitly supposed to be by those who discuss the incidence of aesthetic considerations in theory-evaluation, the historical development of science surely escapes the explanations proposed for it by the rationalist image.

The present work aims to establish that, while both these lower-level models of scientific methodology are to a large measure valid, their validity does not compel the abandonment of the rationalist image as the top-level model in the hierarchy. Indeed, this work will attempt to show how under the rationalist image of science can be accommodated both the belief that the methodological canons of science have undergone radical and sudden transformations, and the belief that aesthetic considerations lie among the grounds on which scientific communities have chosen between competing theories. To this end it will be necessary to contribute to the development of the two lower-level models which initially appear incompatible with the rationalist image of science. Among the tasks of the present work are thus the elucidation of the concept of the methodological revolution in science and the study of scientists' practice of aesthetic theory-evaluation. Furthermore, since there is a conceptual interdependence between models at lower levels of the hierarchy and the top-level model which claims consistency with them, the very act of contributing to the development of the lower-level models will tend to alter the model at the top. The present work will thus lend a certain new gloss to the rationalist image: its outcome will consist of elements of a rationalist model of science which nonetheless allows the occurrence of revolutionary changes in methodological canons and reserves a distinctive role to aesthetic considerations in theory-choice.

It was remarked above that the models of scientific revolution currently circulating in the philosophical community have attained a degree of sophistication considerably greater than that of any available model of the practice of aesthetic theory-evaluation in science. The latter practice thus calls for
more study than does the former. Furthermore, the alterations which the present treatment intends to impose to the commonly-accepted model of scientific revolution derive from the model which it will construct of the practice of aesthetic theory-evaluation. In consequence, the bulk of the following treatment will be devoted to a construction of a model of the practice of theory-evaluation on aesthetic grounds in science; the incidence of this model upon notions of revolutionary discontinuities in science will be studied in chapter 8, and the final chapter will show the relevance of these conclusions to the rationalist image.

Before this enterprise is joined, it is necessary to set out some of the methodological presuppositions by which it will be guided. This will be achieved in the remainder of the present chapter.

2. Mutual evaluations of science and philosophy

History and philosophy constituted according to logical positivism complementary but quite separate approaches to the study of science: on this view the historian sought to construct an accurate and well-supported description of actual scientific practice whilst the philosopher issued a normative, evaluational and largely a priori account of how science ought to proceed. A pecking-order of the disciplines was thus erected. Philosophers occupied the prescriptive high ground from which to test and evaluate scientists' beliefs and methodologies, as these had been described by historians; but philosophical theories concerning science were believed subject to no countervailing evaluation, since philosophy – it was argued – was not an empirical discipline and did not therefore depend upon the accuracy of factual beliefs about science. The extent to which actual science departed from the philosophers' rational reconstructions was interpreted as a measure not of the misunderstanding of science by philosophers but on the

---

3 The current need for further study of the practice of aesthetic theory-evaluation is sensed for instance by Barrow in a review of a book of Dyson's which describes the appeals by Dirac and other scientists who are - in the author's expression - 'unifiers' to considerations of beauty in theory-assessment. Barrow finds Dyson's remarks unsatisfying (1988, p. 1,392): 'Missing from the argument [...] is [...] why there is a peculiar tendency for scientists who are unifiers to hypostatize the notion of "beauty" in their mathematical equations. Is Dirac saying anything sensible about aesthetics when he argues that it is of paramount importance to have "beauty" in one's equations, or is he merely revealing a rather limited personal experience and appreciation of things other than equations?' In chapter 5 and elsewhere below, a rationale for Dirac's appeals to beauty in the appraisal of theories will be offered.
CHAPTER ONE: THEMES AND PROCEDURES

contrary of the shortcomings of real, imperfect and messy science by comparison with its ideal form of which the features were philosophically stipulated.

Of course this view has in recent years been heavily criticized by among others T.S. Kuhn, I. Lakatos and L. Laudan. Whilst granting that the aim of the philosophy of science is in part the generation of a set of norms by which to evaluate scientific theories, such critics envisage also a countervailing evaluation of philosophical theories in the light of scientific practice. After all, argues Laudan, we possess an intuitive grasp – bequeathed by a low-level acquaintance with the history of science – of what constitutes paradigmatically successful science, and a philosophical reconstruction of the discipline should render justice to this apprehension.4 Philosophical systems are consequently subject to judgements based on historical data. We should reject an account of for instance the methodology of theory-assessment which construed as unscientific many acts of theory-choice in periods of the history of physics or biology in which, we intuitively believe, those disciplines were scientifically successful.

Philosophy and science thus lie in an evaluative closed circle. On the one hand philosophy of science holds a normative import for scientific practice and can therefore evaluate competing scientific theories by reference to their philosophical features (e.g. their logical structure or their use of conceptual elements such as hypotheses or empirical data); on the other hand data from the historical narrative of scientific practice exercises an evidential role in philosophy of science and may militate for or against competing philosophical theories about science by supporting or undermining their claim to depict actual science.5

The philosophy of science is by no means the sole branch of philosophical enquiry in which theories of a prima facie normative nature are evaluated by reference to data from an empirical source: such an evidential relation is encountered whenever philosophical speculation attempts to reconstruct an extant practice. An empirically-based evaluation of a philosophical theory similar to those performed in the philosophy of science is met by moral or political philosophers in, say, initial attempts to elucidate the notion of ‘justice’. In the Republic of Plato, the first definition of ‘justice’ proposed by Cephalus is ‘telling

---

5 For a further discussion of these issues see McAllister (1986), pp. 315-8. That philosophical reconstructions ought to bear some resemblance to historical accounts of science is conceded even by Giere (1973), who distinguishes himself in the discussion of this problem by his wish to deny any relevance of historical data to the philosophy of science.
contrary of the shortcomings of real, imperfect and messy science by comparison with its ideal form of which the features were philosophically stipulated.

Of course this view has in recent years been heavily criticized by among others T.S. Kuhn, I. Lakatos and L. Laudan. Whilst granting that the aim of the philosophy of science is in part the generation of a set of norms by which to evaluate scientific theories, such critics envisage also a countervailing evaluation of philosophical theories in the light of scientific practice. After all, argues Laudan, we possess an intuitive grasp – bequeathed by a low-level acquaintance with the history of science – of what constitutes paradigmatically successful science, and a philosophical reconstruction of the discipline should render justice to this apprehension. Philosophical systems are consequently subject to judgements based on historical data. We should reject an account of for instance the methodology of theory-assessment which construed as unscientific many acts of theory-choice in periods of the history of physics or biology in which, we intuitively believe, those disciplines were scientifically successful.

Philosophy and science thus lie in an evaluative closed circle. On the one hand philosophy of science holds a normative import for scientific practice and can therefore evaluate competing scientific theories by reference to their philosophical features (e.g. their logical structure or their use of conceptual elements such as hypotheses or empirical data); on the other hand data from the historical narrative of scientific practice exercises an evidential role in philosophy of science and may militate for or against competing philosophical theories about science by supporting or undermining their claim to depict actual science.

The philosophy of science is by no means the sole branch of philosophical enquiry in which theories of a prima facie normative nature are evaluated by reference to data from an empirical source: such an evidential relation is encountered whenever philosophical speculation attempts to reconstruct an extant practice. An empirically-based evaluation of a philosophical theory similar to those performed in the philosophy of science is met by moral or political philosophers in, say, initial attempts to elucidate the notion of ‘justice’. In the Republic of Plato, the first definition of ‘justice’ proposed by Cephalus is ‘telling

---

5 For a further discussion of these issues see McAllister (1986), pp. 315-8. That philosophical reconstructions ought to bear some resemblance to historical accounts of science is conceded even by Giere (1973), who distinguishes himself in the discussion of this problem by his wish to deny any relevance of historical data to the philosophy of science.
the truth and paying back anything we may have received.' Socrates criticizes this suggestion as inadequate, putting forth an example of an act which falls within the proposed definition but which, he alleges, is ruled to be 'unjust' by the intuitive notion of 'justice' shared by all the interlocutors: if a friend who left with me a weapon has since gone mad, the act of returning the weapon to its owner, whilst in accordance with the proposed definition of Cephalus, would be unjust. Since candidate-definitions of 'justice' are ostensibly prescriptive, aiming at governing human conduct, it is a matter of some initial mystery why a proposed version of that definition should be vulnerable to the observation that a certain human action would fall within or without it. The answer is that attempts to define 'justice' aim in part to account for existing practice, to which Socrates here refers.

Criticism of this kind establishes alongside the logicist analysis of science performed by philosophy tout simple (most brilliantly by the logical positivist school) a historicist reconstruction erected by a compound discipline of history and philosophy of science, or HPS. The difference between these two approaches to the study of science lies in the role reserved therein to historiographic evidence. The only role open to historical data in purely philosophical analyses of science is that of exemplification: philosophical disputes cannot, presumably, be decided by data on what as a matter of fact transpired in historical episodes. Issues raised in HPS are on the contrary in principle susceptible of decision by historical evidence.

The pursuit of the conjoined discipline of HPS appears at first blush to initiate an infinite judgemental regress. Under the tenets of logical positivism, the task of selecting a philosophical standpoint from which to assess the relative merits of scientific theories was conceptually independent of any features of those or of any other actual theories and could therefore in principle be completed before any investigation was undertaken of the vicissitudes of actual science; in other terms, it was logically possible to perfect one's philosophy of science before embarking on the study of its practice or history. In the light of the more recent criticism such a project appears nonsensical: the construction of a philosophy of science has now come to depend to a certain extent upon an empirical input from the history of science concerning the features of successful or representative

---

6 Republic, 331c (Cornford ed. 1968, p. 7). My attention was drawn to this passage by H.I. Brown (1977, p. 133), who however uses the illustration to different effect.

7 Ibid. In the face of this criticism Cephalus withdraws the definition which he proposed.
scientific theories. Here is initiated the regress: for in turn how is one to assess which such theories are indeed representative or successful – and thus warrant consideration in the process of constructing one's philosophy of science – without a prior philosophy from the standpoint of which to perform the judgement?

The circle is not as vicious as may at first glance appear. The solution is plainly that history and philosophy of science are now to be considered no longer separate disciplines but rather a single unified field which accomplishes its evaluations by a recursive effort in which subsequent assessments of philosophical and scientific theories are mutually corrected and perfected. The question does not arise of which of the two disciplines is logically prior and is therefore subject to being completed without recourse to the other, since we have here not two but a single disciplinary approach to science.

The work which follows locates itself within the conjoined discipline of HPS described in this section rather than in the tradition of logicist investigations of scientific method. As a consequence, historical evidence is deemed in this work to possess an evidential and evaluative role rather than to be a mere source of exemplification. This stipulation helps to explain why the treatment should devote so much space to material culled from historiographic sources whilst being in principle a work in philosophy of science: the historical evidence marshalled in chapter 10 and elsewhere has in this treatment the power of contributing to decide the degree of validity to be attributed to the philosophical theses contained herein.

3. The value of scientists' testimony

In tracing the progressive evaluative circle of which the previous section speaks, historically-informed philosophers of science have customarily looked to so-called rational reconstructions as the source of historiographic tests of their conclusions. A rational reconstruction is an account of some – preferably – commonly-known and non-controversial historical episode which makes crucial use of the conceptual apparatus of the philosophical thesis under test. The intention is that the degree to which the rational reconstruction accords with what is known about

---

8 A survey of the debate on the relationship between philosophy of science and historical evidence is contained in Losee (1987), passim.
CHAPTER ONE: THEMES AND PROCEDURES

the episode independently of the claims of the thesis – and this includes what is known 'pre-analytically' about the episode – should be translated into empirical support for the philosophical thesis.9

On the other hand, philosophers of science have tended to recoil from enlisting in support of their views about the manner in which science has progressed the statements on this topic of scientists themselves. Even if a certain scientist's description of the manner in which a piece of research had been undertaken or an inference drawn happened to accord with the philosophically-informed description of this episode proposed by a given thesis in philosophy of science, in the view of many philosophers that testimony cannot be allowed to count as evidence in support of the thesis. The reason for ruling such prima facie evidence as inadmissible is the fear that scientists may constitute unreliable witnesses about their own and their colleagues' actions: to study scientific method, the motto runs, one ought to examine what scientists do rather than what they say that they do.10

The present treatment recognizes the sagacity of scrutinizing the degree of reliability of items of historiographic evidence which are adduced in support of a thesis in philosophy of science from whatever source these items may derive; but it does not entirely share the special and further fears about the use in evidence of scientists' testimony which are outlined in the previous paragraph. In fact, this treatment will at several junctures appeal to scientists' testimony – as well as to the rational reconstructions contained in chapter 10 – in support of its claims about scientific method. Since this practice perhaps departs from the usual mode of argument in the discipline, some remarks in its defence are required. Two justifications for a limited appeal to scientists' testimony in support of philosophical theses are here envisaged.

First, the evaluative circle which was described in the previous section and which constitutes a common mode of progress in the discipline of HPS is to a certain extent susceptible to the danger of excessive conservatism. Consider the operation of the evaluative circle in a particular hypothetical scenario. If the discipline of HPS were to witness a run of philosophical theories about scientific method which did not admit the occurrence of a certain class of phenomena in

---

9 Some further discussion of rational reconstructions will be pursued in the following chapter.

10 Some of the doubts on the reliability of scientists' pronouncements are motivated by the growing understanding of their rhetorical purposes: representative studies of this aspect of the methodological statements of scientists are contained in Schuster and Yeo eds (1986).
CHAPTER ONE: THEMES AND PROCEDURES

scientific practice, it is likely that an increasing proportion of histories of science would then come to be written which, influenced by the philosophical consensus, would suppress or fail to perceive the historiographic evidence which suggested the occurrence of phenomena of that class. A growing degree of eminence would most likely come to be attached within the discipline to the historiographic model encapsulated in those treatises which omits mention of that particular class of phenomena of scientific practice. Then the rational reconstructions of episodes of history to which philosophers appealed in support of their claims would in turn tend to fall under the influence of what would be seen as the historiographic consensus about past science: the continuing development of the philosophy of science would further be insulated from evidence of the occurrence of the phenomena which had previously been ignored.

While this tendency in the pursuit of the discipline of HPS does not amount to a strict vicious circle – in that a new historiographic finding or an innovation in philosophical thinking can at any moment recall the attention of the community upon the previously-neglected class of phenomena – it can needlessly and damagingly limit the scope of studies in the discipline. This limitation is evaded most easily by appealing to a body of historiographic evidence which lies outside the evaluative circle established in the discipline, which has not already been stripped of unfamiliar suggestions by the operation of the circle, and which may therefore force upon the awareness of the community a new range of phenomena to be investigated. Appeal to historiographic evidence in forms other than the currently popular rational reconstructions can therefore alert the community to the occurrence of previously unsuspected or neglected phenomena in scientific practice. Among the items of historiographic evidence which can exercise this role is, notably, the reported testimony of scientists.

When used to this end, scientists’ testimony is a body of evidence apt to suggest that the scope of philosophical or historical enquiry about science be widened to embrace the study of some aspect of or phenomenon in scientific practice which has been neglected by the evaluative circle composed of successive rational reconstructions. The suitability of scientists’ testimony to attain this end constitutes a powerful justification for its use by studies in HPS. Once the scope of the inquiry has been broadened on the suggestion of scientists’ testimony, the philosopher may return to customary methods for constructing theories about science, and appeal to rational reconstructions of episodes of history to gauge the validity of the theories which result. Since in this usage scientists’ testimony plays primarily a suggestive rather than validational role, any misrepresentation of
scientific practice which it may contain is not necessarily reflected in the theories embraced about science by philosophers, if these are properly tested by appeal to rational reconstructions.

The present work will deploy historiographic data in the form of scientists' testimony as evidence for the claim that the appeal in certain contexts to aesthetic considerations form a customary part of scientific practice, and are hence a proper object of study by the philosopher. The reason for appealing to scientists' methodological pronouncements to this end is the wish to break the tacit consensus in current writing in the philosophy of science that aesthetic considerations are of precious little interest in the construction of a rational image of science. Where aesthetic criteria have been mentioned in recent work in philosophy of science, their incidence has generally been confined to the domain of theory-formulation, as will be shown in chapter 5; in any case, the criteria to which a scientist has appealed have generally been supposed to have been peculiar to him or her and not the expression of a methodological norm entrenched in a social practice. In consequence, aesthetic considerations have generally been considered unimportant to the elucidation of the nature of scientific rationality.

Under the operation of an evaluative circle of the kind mentioned above, current influential works in the historiography of science have conformed to this tacit philosophical consensus: today historians' mention of aesthetic predispositions is confined to biographical history, playing little or no part in the project of reconstructing the large-scale behaviour of scientific communities in, say, instances of theory-choice. In the face of the seeming consensus on the part of historians and philosophers to ignore as irrelevant the aesthetic predispositions of scientists in theory-choice, appeal to scientists' testimony directly rather than through the mediation of professional historians may serve to alert historians and philosophers to the need to broaden the scope of their models of science and reserve a place of greater importance than has hitherto been common to the study of the aesthetic considerations which figure in scientific practice.

The second justification for the appeal to scientists' testimony which is envisaged here has less force, and seeks to suggest that this form of historiographic evidence is prima facie no less worthy of consideration than are historiographic data accessed in the form of a rational reconstruction.

Section 1 of this chapter suggested that the attempt to reduce scientific methodology to comprehension is pursued by erecting a hierarchy of partial models of science, each of which accounts for some aspect of scientific practice
and the conjunction of which possesses some high degree of internal consistency. Rational reconstructions of episodes of the history of science can most aptly take their place in this hierarchy, since such reconstructions are just models of the episodes to which they are associated. But the methodological pronouncement of a scientist, which contains the scientist’s own description of his or her actions at a historical juncture, is similarly a partial model of the history of science. For sure, the scope of the model constituted by the rational reconstruction may be much broader than that of the model constituted by the scientist’s memoir, in that the former may purport to account for each member of a large set of historical episodes while the scientist may have uttered a statement intended to pertain to no more than a singular event; but the scenario envisaged in section 1 is that of a hierarchy of models of varying scope, so a hierarchy in which some of the models have a scope broader or even much broader than that of others is just what one would hope to assemble. The conclusion which this comparison supports is that rational reconstructions of historical episodes on the one hand and the descriptions of scientific methodology uttered by scientists on the other are eiusdem generis in being bodies of historiographic data to which an initial theoretical interpretation has been lent. Provided that the cautions which are used in approaching rational reconstructions are used also in the reading of scientists’ testimony, the latter is in principle a no less useful source of historiographic data than are the former.11

11 For further discussion of the value of scientists’ methodological testimony see Kragh (1987), pp. 150-8.
Chapter Two
THE STIPULATION OF METHODOLOGICAL NORMS

1. War games and the philosophy of science

Many who are not otherwise militaristic turn for recreation to war games, in which model armies are commanded into battle on gameboards by contestants subject to regulations. The rules lend structure to the game by setting constraints upon the range of actions permitted of the players, by imposing upon them a tactical canon. This canon may be explicitly demanded, through the prescription to the players of specific deliberative formulae, or implicitly promoted, perhaps by suitably defining the imagined performance of each class of model soldiers at their command.

In principle the tactical canon and the consequent character of a war game are wholly arbitrary, and one may invent games of any complexion by adjusting the rules. With greater representationalism, however, some war games aim not to stage an imaginary conflict but to replay a battle of recorded history. Here the players enter the boots of historical military commanders in a reconstruction of the outbreak of a past battle which they thereafter manipulate in accordance with a tactical canon of some degree of verisimilitude. Both the initial conditions and the rules of such a Kriegsspiel are informed by historical data on the actual battle: these may relate to the topography of the battleground, the numbers and initial disposition of the opposing forces, and the specifications of their weapons.\footnote{In a typical war-gamers' manual, Featherstone (1977), pp. 19-92 describe the forces which participated in various tank-battles in the Mediterranean theatre of the Second World War, the terrain which they encountered and the tactics which they pursued, pp. 93-114 draw from the outcomes of those battles lessons for the tactics of those who would re-enact them in a war game, and pp. 115-46 list the specifications of the vehicles and weapons employed therein. These particulars are intended to help to mount a sufficiently realistic war game and draft its rules of play. There circulate innumerable primers of this kind for what is evidently a widespread pastime. Such manuals are interesting historiographical beasts: falling short of treatises of military history of the Thucydidean model, they limit themselves instead to providing the 'boundary conditions' of episodes of actual history from which re-enactments of those episodes will originate and within the confines of which they will develop.}
CHAPTER TWO: THE STIPULATION OF METHODOLOGICAL NORMS

players may yet exercise freedom of action, and their decisions in the course of the game may learn and diverge from the conduct of the historical commanders to the extent of reversing the outcome of their encounter. The scope of the analogical relation is therefore limited: although the gameboard may map the field of combat and the ranks of figurines may constitute a model of the historical armies, the re-enacted engagement does not necessarily model the historical episode. A game may open on a reconstruction of the field of Marengo but in its progress assume configurations quite different from those of the battle of 1800.

One may however conceive of extending the scope of the model by altering the tactical canon imposed upon the players by the rules. By judicious redrafting of the rules it is in principle possible to replicate in the players of the game the string of decisions followed by the commanders in the actual battle which the game re-enacts. From the historical initial conditions of a battle will ensue under these rules its historical dénouement: re-enactments of Marengo will see Napoleon for evermore defeat the Austrians. The scope of the model now encompasses not just the physical manifestations of the battle but also the conditions which prompted the inferences and decisions of the historical participants. The entity which models those conditions is the war game’s rule-book. For the modelling of the historical combatants’ sequence of decisions to achieve success, the combatants need not have explicitly endorsed the tactical canon compiled for the players of the subsequent game; a historian may however find evidence separate from the combatants’ sequence of decisions – contained in for instance their descriptions of their own actions – that the combatants acknowledged the concerns or objectives set by the game’s rule-book.

Part of the discipline of philosophy of science stands in a relation to the historical record analogous to that of a war game. As the war game is a study of the judgements and decisions of a historical military commander, conducted through a re-enactment of the scenario and constraints encountered by that figure, so the philosophy of science studies the judgements and inferences of the historical scientist through a reconstruction of the situation and constraints faced by that figure. As the war-game player first drafts a proposed rule-book, then investigates how a personator of Napoleon acts at a reconstruction of Marengo under the constraints of that book, and finally draws from the similarities and divergences between the re-enactment of the battle and the historical Marengo conclusions concerning the capacities and constraints of Napoleon, so acts the philosopher in regard to the historical scientist. Under this analogical relation the relatum of the war-game rule-book is the system in prescriptive philosophy of
players may yet exercise freedom of action, and their decisions in the course of the game may learn and diverge from the conduct of the historical commanders to the extent of reversing the outcome of their encounter. The scope of the analogical relation is therefore limited: although the gameboard may map the field of combat and the ranks of figurines may constitute a model of the historical armies, the re-enacted engagement does not necessarily model the historical episode. A game may open on a reconstruction of the field of Marengo but in its progress assume configurations quite different from those of the battle of 1800.

One may however conceive of extending the scope of the model by altering the tactical canon imposed upon the players by the rules. By judicious redrafting of the rules it is in principle possible to replicate in the players of the game the string of decisions followed by the commanders in the actual battle which the game re-enacts. From the historical initial conditions of a battle will ensue under these rules its historical dénouement: re-enactments of Marengo will see Napoleon for evermore defeat the Austrians. The scope of the model now encompasses not just the physical manifestations of the battle but also the conditions which prompted the inferences and decisions of the historical participants. The entity which models those conditions is the war game’s rule-book. For the modelling of the historical combatants’ sequence of decisions to achieve success, the combatants need not have explicitly endorsed the tactical canon compiled for the players of the subsequent game; a historian may however find evidence separate from the combatants’ sequence of decisions – contained in for instance their descriptions of their own actions – that the combatants acknowledged the concerns or objectives set by the game’s rule-book.

Part of the discipline of philosophy of science stands in a relation to the historical record analogous to that of a war game. As the war game is a study of the judgements and decisions of a historical military commander, conducted through a re-enactment of the scenario and constraints encountered by that figure, so the philosophy of science studies the judgements and inferences of the historical scientist through a reconstruction of the situation and constraints faced by that figure. As the war-game player first drafts a proposed rule-book, then investigates how a personator of Napoleon acts at a reconstruction of Marengo under the constraints of that book, and finally draws from the similarities and divergences between the re-enactment of the battle and the historical Marengo conclusions concerning the capacities and constraints of Napoleon, so acts the philosopher in regard to the historical scientist. Under this analogical relation the relatums of the war-game rule-book is the system in prescriptive philosophy of
science, battle-tested for its effects on the evolution of the re-enactment of the historical episode; the \textit{relatum} of the tactical canon prescribed of the war-game player by the rules is the canon of rationality tentatively ascribed to the historical scientist; and the \textit{relatum} of the commander whose behaviour is thus studied is the scientist of interest.

Such re-enactments are performed whenever a body of modern methodological precepts is brought to bear upon a scenario encountered by past science, and it is investigated how the protagonists of the period would have acted if following those dictates. A historical figure whose research and inferences have been re-enacted under the aegis of several competing philosophical systems is of course Galileo: one meets the intellectual biographies of Galileo as a Platonist, an Aristotelian, an inductivist and an experimentalist.\textsuperscript{2} The accounts therein contained of key junctures of Galileo's research are reconstructions which proceed along lines dictated or constrained by the philosophical rule-books to which they pay allegiance.

In principle the fidelity of these scientific re-enactments is no less flexible than that of their military counterparts. Under rules which allow sufficient freedom of action, the players may depart from the inferences and decisions of the original participants of the historical episodes now under reconstruction: indeed a philosophical system may permit the personators of the historical scientists to learn and benefit from the errors of their archetypes and thus to hasten in the re-enactment the advancement of science. This degree of freedom is allowed for instance by the philosophical system of Lakatos. One of the introductory statements in his treatment of historiography signals the congruity between his vision of the philosophy of science, which assigns a central role to reconstructions of past scientific episodes, and the present treatment based on the analogy with war games: Whatever problem the historian of science wishes to solve, he has first to reconstruct the relevant section of the growth of objective knowledge, that is, the relevant section of "internal history".\textsuperscript{3} Lakatos allows the historian bent upon the reconstruction of this section of history a wide freedom of operation:

\begin{quote}
In constructing internal history the historian will be highly selective: he will omit everything that is irrational in the light of his rationality theory. But this normative selection still does not add up to a fully
\end{quote}

\textsuperscript{2} These interpretations have been propounded respectively by Koyrè (1939), Wallace (1961), Whewell (1840) and Drake (1970). The character named Galileo who features in certain of the more partisan interpretations (less sagacious than those mentioned here) is, some readers would say, clearly an impersonator.

\textsuperscript{3} Lakatos (1971), p. 106.
science, battle-tested for its effects on the evolution of the re-enactment of the historical episode; the *relatum* of the tactical canon prescribed of the war-game player by the rules is the canon of rationality tentatively ascribed to the historical scientist; and the *relatum* of the commander whose behaviour is thus studied is the scientist of interest.

Such re-enactments are performed whenever a body of modern methodological precepts is brought to bear upon a scenario encountered by past science, and it is investigated how the protagonists of the period would have acted if following those dictates. A historical figure whose research and inferences have been re-enacted under the aegis of several competing philosophical systems is of course Galileo: one meets the intellectual biographies of Galileo as a Platonist, an Aristotelian, an inductivist and an experimentalist.\(^2\) The accounts therein contained of key junctures of Galileo's research are reconstructions which proceed along lines dictated or constrained by the philosophical rule-books to which they pay allegiance.

In principle the fidelity of these scientific re-enactments is no less flexible than that of their military counterparts. Under rules which allow sufficient freedom of action, the players may depart from the inferences and decisions of the original participants of the historical episodes now under reconstruction: indeed a philosophical system may permit the personators of the historical scientists to learn and benefit from the errors of their archetypes and thus to hasten in the re-enactment the advancement of science. This degree of freedom is allowed for instance by the philosophical system of Lakatos. One of the introductory statements in his treatment of historiography signals the congruity between his vision of the philosophy of science, which assigns a central role to reconstructions of past scientific episodes, and the present treatment based on the analogy with war games: 'Whatever problem the historian of science wishes to solve, he has first to reconstruct the relevant section of the growth of objective knowledge, that is, the relevant section of "internal history".'\(^3\) Lakatos allows the historian bent upon the reconstruction of this section of history a wide freedom of operation:

In constructing internal history the historian will be highly selective: he will omit everything that is irrational in the light of his rationality theory. But this normative selection still does not add up to a fully

\(^2\) These interpretations have been propounded respectively by Koyré (1939), Wallace (1981), Whewell (1840) and Drake (1970). The character named Galileo who features in certain of the more partisan interpretations (less sagacious than those mentioned here) is, some readers would say, clearly an impersonator.

\(^3\) Lakatos (1971), p. 106.
fledged rational reconstruction. [...] Internal history is not just a selection of methodologically interpreted facts: it may be, on occasion, their radically improved version. 4

Those who in a Lakatosian reconstruction re-fight a past scientific battle are in truth fighting a battle quite different from the original, with a new tactical canon, dissimilar moves and a discrepant outcome. Lakatos quite openly accepts this divergence:

One may illustrate this using the Bohrian programme. Bohr, in 1913, may not have even thought of the possibility of electron spin. He had more than enough on his hands without the spin. Nevertheless, the historian, describing with hindsight the Bohrian programme, should include electron spin in it, since electron spin fits naturally in the original outline of the programme. Bohr might have referred to it in 1913. 5

Clearly whoever will personate Bohr in a Lakatosian reconstruction is not constrained to follow Bohr's original inferences, any more than the war-gamer who takes the part of Napoleon at Marengo is necessarily compelled to deploy his or her model troops in the same actions. Both are free to improve upon the performance of their historical archetypes. 6

In advocating the pursuit of reconstructions in which modern impersonators of past scientists are allowed to learn from and thus correct the blunders of their predecessors, Lakatos makes a virtue out of historiographical inexactitude. He severely reduces, however, the specificity of the historian's work. A figure in Lakatos's history will resemble an actor who, while remaining in character, departs from the prescribed lines to utter ad-lib comments to the audience: though this knack may please in its inventiveness, it reveals little of the texts on which the performance is based. The greater historiographic insight is naturally obtained when the re-enactment is capable of capturing the vicissitudes of the original episode. By careful tuning of the variables which affect the behaviour of the reconstruction, the battles relived by philosophers of science can

4 Ibid.; emphasis in the original.
5 Ibid., pp. 106-7. Lakatos notoriously continues: 'Why Bohr did not do so, is an interesting problem which deserves to be indicated in a footnote' (ibid., p. 107).
6 Musgrave (1976, p. 193) openly concedes that the Lakatosian 'rational reconstruction' in which he is engaged amounts to 'remedying history's mistakes'. Here and there in Howson ed. (1976), a collection of historical essays on Lakatosian principles which includes Musgrave's paper cited here, one finds evidence that the authors regard their studies as a re-staging of the episodes, as investigations of how the protagonists would have acted if endowed with the canons of rationality which Lakatos's philosophical system predicates of the characters in the reconstruction. Musgrave again for instance writes (1976, ibid.): 'In rationally reconstructed history, having predicted oxygen, Lavoisier went on to discover it – while in actual history it was discovered by Priestley.'
like those re-enacted by military enthusiasts come to reproduce the outcomes of their historical correspondents. The rules then possess both a normative and a descriptive valence: they shape the turns taken by reconstructions of historical episodes and, to the extent to which those reconstructions are faithful to the historical evidence, they model the conditions under which the participants in those episodes drew their inferences and took their decisions. The scope of the model constituted by the philosophical system is then extended and embraces the most recondite and impalpable element of the past episode: the canon of rationality proposed by the philosophical system comes to model the actual canon which animated the characters of the past.

Once again, as in war games, for the modelling of past sequences of scientific inferences to achieve success, the authors of those inferences need not have endorsed or even recognized the norms or priorities enshrined in the ‘rules of play’ prescribed by philosophical systems to their later impersonators; nonetheless scientists frequently describe the methodological values which they consider have animated their work, and those stated values may accord with those prescribed in the reconstruction, thus further enhancing the fidelity of the model. Indeed a reading of the methodological reflections of past scientists may well prompt the inclusion of the values which they mention in the rule-books of later philosophical systems.

The analogy of war-gamers’ rule-books is able to shed intuitive light on the relationship of philosophical systems to historical evidence. As the stakes of the philosophical battles are however of greater conceptual subtlety than those of military re-enactments, the debate on the compilation of the historiographic rule-book has been correspondingly more intense.

---

7 The conception of the function of rules in the philosophy of science which is prompted by the analogy with war-game rule-books may help to reconcile Feyerabend to prescriptivism in this arena. Feyerabend (1975, e.g. p. 23) mistrusts the prescription of rules to scientists, fearing it will reduce their methodological freedom and inhibit progress. But in the present treatment, which might be termed one of vicarious prescriptivism, the objects of such prescriptions are not living scientists but imaginary agents delegated to re-enact an episode of scientific history. Since the creativity of living scientists is thereby unaffected, Feyerabend would surely have no grounds to resist this use of methodological precepts.
CHAPTER TWO: THE STIPULATION OF METHODOLOGICAL NORMS

2. The constitution of scientific success

If a purpose of the rule-book in a war game is to achieve in the re-enactment a close replica of the original episode, any proposed version of the rule-book may seek to draw justification from the closeness of the replication. This historiographic warrant is not of course the sole form of justification which a rule-book may seek to gather: there exists a class of justificatory argumentation which is non-historiographic or systematic. In the case of war-game rule-books this form of justification is probably rather trivial, and derives from common-sense understanding of the goals and nature of warfare: nonetheless, because of the existence of this second source of justification, any proposed rule-book will be evaluated not solely on its capacity to reproduce on the gameboard the vicissitudes of past battles, but also on its acceptability on common-sense grounds.

Systematic justification plays a role of much greater moment in the other pole of the extended metaphor here being explored, in the compilation of methodological rules in the philosophy of science. The norms which are prescribed to the personators of past scientists in a reconstruction are justified not solely by the accuracy of the resulting reconstruction but also by philosophical argumentation from the goals and nature of science. These arguments originate in a judgement of the identity of the theoretical quality to be held constitutive of success in science, and aim to determine a canon of methodological precepts apt to generate theories which exhibit that quality to the highest degree.

The identification of what counts as scientific success seems at first blush afflicted by a relativism with respect to ontological doctrine: while antirealism, shunning ontological commitment, views success as the attainment of empirical adequacy, realists posit as the ultimate goal of science the construction of true explanatory theories about phenomena. The divergence in ontology is however recomposed in methodology. Antirealists readily recognize – indeed proclaim – observational success as the sole possible content of the notion of empirical adequacy, while realists for their part see observational success as the sole – albeit a fallible – diagnostic criterion of truth or verisimilitude in the absence of any means of comparison with a putative ‘truth of the matter’ to establish how closely

---

8 Losee (1987, pp. 77-110) similarly distinguishes between what he terms ‘historicist’ and ‘logicist’ justifications of scientific methodologies.

9 The stipulation of explanatory power eliminates from the quest tabulations of mere logical theorems or phenomenal accidents and introduces a requirement for theories to grasp the causal mechanisms which underlie reality.
a theory has fulfilled the goal of true explanation. Observational success is thus consensually and unproblematically held to be constitutive of what it is for a scientific theory to be successful.\(^\text{10}\)

Ultimately, a theory's observational success would be demonstrated by its according with evidence accumulated from disparate sources over unlimited time-spans. But unlimited time-spans are not at the disposal of scientists hesitating between alternative theories: the working scientist requires criteria of prompt application able to indicate whether a theory is likely to demonstrate long-term observational success. A set of such criteria is constructed tacitly by a scientific community and explicitly by philosophers of science by considering what features it would be necessary or helpful for a theory to possess for it to be able to demonstrate long-term observational success. Since observational success would be a logico-empirical attainment of a theory, the criteria to which these considerations lead are themselves logico-empirical in nature: that is, they judge the theory to which they are applied on its possession of certain logico-empirical features, an example of some of which is offered in the next section.\(^\text{11}\)

In this way is constructed a chain of reasoning from the stated goal of science, observational success, to a set of logico-empirical features of theories. The chain of reasoning yields a conditional imperative: if contributions are to be made towards the attainment of the goal of science, theory-choice must prize those theories which exhibit a certain set of logico-empirical features.

A chain of inference from the goal of science to a canon of methodological precepts lies at the heart of the model of scientific methodology of W.H. Newton-Smith. He first states the aim which he attributes to science: 'One of our starting points in this work was the assumption that the goal of science is truth.'\(^\text{12}\) His construal of the notion of truth quickly leads to the formulation of a eventual principle of comparative theory-test: 'In the long run the ultimate test as to whether one theory has more successfully latched on to a facet of the world than another theory is their relative observational success.'\(^\text{13}\) The inordinate length of the time-spans over which for this test to be reliable it is necessary to allow a

\(^{10}\) The wide agreement on this point throughout the philosophical community is one of the reasons for which the debate between realists and antirealists is so apparently undecidable: if realists and antirealists differed in the substantive methodological prescriptions which they issued to scientists, it might be possible to conclude that one or the other ontological doctrine was superior on those grounds.

\(^{11}\) That logico-empirical criteria have attracted the lion's share of the attention in philosophical discussions of the appraisal of scientific theories is attested by Koertge (1979), pp. 228-40, in a review of current research in the field.


\(^{13}\) Ibid.
CHAPTER TWO: THE STIPULATION OF METHODOLOGICAL NORMS

to accumulate observational success then prompts Newton-Smith to infer methodological precepts of quicker applicability: 'We need other factors to guide us which can serve as fallible indicators of likely long-term observational success.' These he identifies in a battery of principles of theory-assessment similar to the logico-empirical criteria to be outlined next.

3. The logico-empirical rule-book

The qualities of theories which are commonly reputed to be constituent of observational success are the following. To a brief account of the reasons for which theories exhibiting each of these qualities are held to be ceteris paribus preferable to others is appended mention of a historical episode in which the perception of that quality in a theory helped to shape its community's attitude to it.

(a) Internal consistency A theory ought not to harbour internal contradictions for, at least in standard logic, any such incongruity would cause it to entail every proposition of its language, and hence be worthless as a predictive tool. Such a flaw was discerned in the Aristotelian theory of free fall which asserted that heavier bodies fell faster than lighter ones. Galileo envisaged a heavy body attached by a cord to a lighter one and asked how fast the resulting compound would be expected to fall. On one application of Aristotelian theory the light body would retard the heavy one so the velocity of the composite should be less than that of the heavier body alone; on another the compound object is heavier than its heavy component mass and should consequently fall faster. Galileo caps his discussion with a ritual formula of one who has caught his opponent in contradiction: 'from the supposition that the heavier body is moved more swiftly than the less heavy, I conclude that the heavier moves less swiftly.'

(b) Consistency with existing corroborated theory P. Duhem and W.V. Quine have

---

14 Ibid., p. 224.
15 For a derivation of the logical result see e.g. Newton-Smith (1981), p. 229. Smith (1988) argues to the contrary that a community need not withdraw support from a theory on account of its possession of internal inconsistencies, especially if the theory appears susceptible of further and interesting development.
16 Galilei (1638), pp. 66-8.
17 Ibid., p. 67.
argued for a form of epistemological holism in which scientific theories are tested not in isolation but in bodies of belief which include auxiliary theories and background knowledge. In domains of science where this view appears pertinent, theories lose their individuality and the notion of a theory is replaced by that of composite bodies of theories: the requirement of internal consistency of a theory correspondingly broadens into a requirement for the mutual consistency of bodies of theories.

Rutherford's atomic theory of 1911, couched within the paradigm of classical electromagnetism, viewed the atom as a miniature planetary system of which the positive charge was concentrated in the centre and surrounded by electrons. A theorem in electromagnetism stated that a system of charged bodies cannot be in equilibrium at rest under the action of their electrostatic interactions alone; simultaneously another result implied that if Rutherford's electrons were in motion along closed orbits, the centripetal acceleration necessary to maintain those trajectories would cause the charges to radiate, dissipate energy, and thus spiral into the nucleus. Rutherford's model thus appeared to contradict the background theory of electromagnetism. This difficulty was resolved by Bohr's model of the atom of 1913 which confined the motion of electrons within quantum-mechanical orbitals, thus declaring classical electromagnetism inapplicable to atomic theory.

Whilst a theory is minimally required not to contradict other, background theories, it is more strongly valued if it can provide deeper explanations of the generalizations which they express. The statistical mechanics of R. Clausius and C. Boltzmann was valued partly for its consistency with the predictively successful empirical laws of classical thermodynamics, and furthermore for deriving the equations of state of ideal gases – which in the earlier theory had the status of mere phenomenological generalizations – from the principles of conservation of energy and momentum.

(c) High degree of simplicity The requirement that theories be simple is in some ways the most ambiguous of the commonly stipulated theoretical desiderata. In many instances the choice by a scientist of one theory in preference to another on the grounds of simplicity appears to be motivated by subjective concerns not immediately connected with the theories' relative empirical performance. It is however reasonably widely accepted that a high degree of simplicity in a theory

---

18 For a survey of various ways in which simplicity in theory-assessment is regarded, see Hesse (1967), pp. 445-6.

19 This observation, which plays a role of some importance in portions of the argument to follow, will be elucidated in chapter 6, where the content of aesthetic canons of theory-assessment is analyzed.
is a logico-empirical virtue of that theory, in that simpler theories making the same claims as more complex ones are better supported by favourable evidence.\textsuperscript{20} S.F. Barker explains the issue as follows:

If one system is simpler than another [...] then the simpler one 'says more,' it has 'more content,' because it excludes a greater number of possible models; therefore it runs more risk of being contradicted by the evidence. A system which takes a risk yet survives deserves more credit, it earns more credibility, than does a system which survives but says less and thus has taken less risk.\textsuperscript{21}

Many applications of criteria of simplicity in theory-evaluation occur in the biological sciences. In evolutionary theory G.C. Williams uses a principle of parsimony to argue that the principles of group selection and biotic adaptation are less well supported by the data than are the principles of genic selection and organic adaptation.\textsuperscript{22} A similar criterion has been extensively applied in the assessment of theories in phylogenetic systematics, as R. Johnson reports.\textsuperscript{23}

\textit{(d) Consistency with known data} Deductions from the theory should lie in agreement with the results of past observation and experiment. Newton rejected the Cartesian theory of vortices as the foundation of celestial mechanics because it was incapable of accommodating among its implications the known elliptical orbits of the planets.

\textit{(e) Generation of novel predictions} If the sole empirical requirement of a theory were its accordance with data previously gathered, the danger would exist that a favourable assessment might be accorded to a theory constructed \textit{post facto} to account for those data. To guard against that danger methodologists commonly stipulate an additional empirical requirement, that a theory should generate some novel predictions. C. Clavius argued that the Ptolemaic theory of the solar system was likely to be true because 'by the assumption of Eccentric and Epicyclic spheres not only are all the appearances already known accounted for, but also future phenomena are predicted.'\textsuperscript{24} W. Whewell insisted that good hypotheses

\textsuperscript{20} Failure to acknowledge this finding is one of the faults of my previous treatment of aesthetic criteria of theory-assessment, McAllister (1989).

\textsuperscript{21} S.F. Barker (1957), pp. 181-2. This explanation of the value of simplicity in theory-preference would win the agreement of Popper: see e.g. his (1959), pp. 140-2.


\textsuperscript{23} Johnson (1982), pp. 83-8. This article provides also a useful brief overview of the discussions of the formal aspects of simplicity-criteria in philosophy of science (ibid., pp. 79-83). For bibliographic advice on this topic I am grateful to Drs D.J. Kornet of the Institute of Theoretical Biology, University of Leiden.

\textsuperscript{24} Quoted in Blake (1960), p. 34. Clavius was writing in 1602.
CHAPTER TWO: THE STIPULATION OF METHODOLOGICAL NORMS

ought to foretell phenomena which have not yet been observed.\textsuperscript{25}

A theory's predictions can be novel in two ways: they may forecast data which before the formulation of the theory one had not thought to gather, or they may 'retrodict' data already available but not previously thought relevant to the assessment of theories in this domain of phenomena.

J. Kepler's theory of planetary dynamics issued a class of predictions which were novel in the former sense. After nine years' investigation of the orbit of Mars Kepler published his first two laws in the \textit{New Astronomy [...] with Commentaries on the Motions of Mars}: their empirical domain extends however - within the bounds of non-relativistic accuracy - far beyond the fourth planet of the Solar System to all celestial bodies and indeed to all bodies in closed orbits under inverse-square law attractive forces, as Newton's work was to establish.

Novel predictions in the second sense are the sign of an attempted unification of previously disparate branches of science. J.C. Maxwell's unification of optics and electromagnetism was responsible for widespread novel predictions in this sense. A measure of the scientific community's appreciation of this achievement is obtained from a eulogy of Maxwell by M. Planck:

Maxwell [...] created a theory which not only could compete with the well established theories of electricity and magnetism but surpassed them entirely in success. For the criterion of the value of a theory, that it explains quite other phenomena besides those on which it was based, has never been so well satisfied as with Maxwell's theory. Neither Faraday nor Maxwell may have originally considered optics in connection with their consideration of the fundamental laws of electromagnetism. And yet the whole field of optics, which had defied attack from the side of mechanics for more than a hundred years, was at one stroke conquered by Maxwell's Electrodynamic Theory; so much so that since then every optical phenomena can be directly treated as an electromagnetic problem. This must remain for all time one of the greatest triumphs of human intellectual endeavour.\textsuperscript{26}

The generation of novel predictions constitutes an important \textit{desideratum} of scientific theories and completes the review of logico-empirical criteria of theory-evaluation.

Although these criteria may be formulated or classified in different ways, their inclusion in a canon by which to evaluate theories is not a matter of current controversy: both scientists and philosophers of science appear to agree that these criteria figure among scientific communities' standard means to judge the logico-

\textsuperscript{25} Whewell (1840), vol. 2, p. 228; emphasis and orthography as in the original.
\textsuperscript{26} Planck (1931), pp. 57-8.
empirical worth of theories. For instance, F.J. Belinfante prefaces a discussion of hidden-variable theories in quantum mechanics by the statement that 'physicists call a theory satisfactory if (1) it agrees with the experimental facts, (2) it is logically consistent, and (3) it is simple as compared to other explanations.'

Newton-Smith, whose approach is representative of that of many philosophers of science, includes among 'the good-making features of theories' observational nesting, fertility, track record, intra-theory support, internal consistency and simplicity.

4. The recourse to extra-empirical criteria

Nothing which has so far been said suggests that of the two views of methodological norms, the historiographic view with a presentation of which the chapter opened and the systematic view briefly discussed in section 2, one possesses philosophical value in any respect greater than that of the other. While a philosophical image of science may seek to justify its methodological precepts through historiographic evidence or systematic argumentation, in keeping with its temperament and predispositions, there has so far been presented no suggestion that one of these courses is preferable to the other.

Where the historiographic conception proves superior to its systematic counterpart is in the task of stimulating philosophers to consider possible additions to their inventory of methodological precepts, and to avert the risk that the list of precepts may be brief to the point of inadequacy. As was argued in the previous chapter, a purely systematic approach to the tabulation of methodological norms, with no historiographic input, tends to spawn an excessively circumscribed methodological canon and a narrow characterization of scientific rationality. Historiographic research, with its access to a virtually inexhaustible pool of historical data, is apt to goad methodologists into consideration of precepts the incorporation of which into the philosophical image of science will allow both a more faithful reconstruction of past scientific episodes and a broader and more interesting characterization of scientific rationality.

It is today a commonplace that ascription to past scientists of logico-

---

27 Belinfante (1973), p. 3.
empirical criteria of theory-assessment alone cannot suffice to reconstruct salient instances of theory-choice. Alone, the logico-empirical rule-book is inadequate to ensure that the staged reconstruction of episodes of scientific history remain close to the events themselves. Accordingly philosophers of science who value the attainment of such proximity have posited additional categories of evaluative criteria, which cannot be reduced to the logico-empirical factors.

The realization that such categories are fruitful for the philosophy of science has however been a gradual achievement, attained in the teeth of rearguard resistance by positivists who privileged logico-empirical criteria to an inordinate degree and therefore denied efficacy or legitimacy to extra-empirical considerations. Throughout much of the debate, the evaluative criteria of the additional categories envisaged by philosophers of science have been deemed by them to be of metaphysical character.²⁹

The positivist attitude towards extra-empirical criteria is exemplified by the writings of H. Margenau and P. Frank. Margenau dedicates a chapter of his survey of the philosophy of modern physics to what he terms the 'metaphysical requirements on constructs'.³⁰ He attributes to them a quasi-Kantian, regulatory role:

> The methodology of science involves deliverances of sense as well as rules of correspondence, constructs, and principles regulating constructs. Having learned that the latter are not conveyed by sensory data and yet function in guiding experience, we should call them metaphysical principles in the modern sense of the world. Metaphysical principles, thus understood, are an important part of all procedures which ultimately define reality.³¹

Margenau names as metaphysical requirements the following: the requirements of logical fertility, of multiple connections, of permanence and stability, of extensibility of constructs, of causality, and of simplicity and elegance.³²

Frank too draws a distinction between empirical and non-empirical criteria of theory-assessment; in consequence of positivist concerns on the demarcation of

---

²⁹ The problem of determining the contribution to theory-assessment of metaphysical criteria is different from, though doubtless related to, that of charting the metaphysical content of scientific theories, even if the theories are subsequently praised or criticized for that component: theory-assessment on metaphysical criteria involves the appeal to metaphysical tenets exterior to the theories examined. Popper’s early criticism of Darwin’s theory of evolution as metaphysical rather than scientific (Popper 1974, pp. 133-43) was thus not an instance of a metaphysical evaluation but a straightforward logico-empirical one – indeed it is for this reason that Popper’s judgement was so unfavourable.


³¹ Ibid., p. 81.

³² Ibid. The criteria of Margenau are briefly discussed by Losee (1987), p. 56.
CHAPTER TWO: THE STIPULATION OF METHODOLOGICAL NORMS

science from pseudo-science, he terms them 'scientific' and 'extrascientific' factors. According to Frank, the "'scientific' criteria in the narrower sense' are 'agreement with observations and logical consistency', a concise formulation of the logico-empirical evaluative canon explored at greater length in the previous section. The 'extrascientific' criteria appear to be ethical and metaphysical; in defence of the inclusion of the former category Frank says 'It is easily seen from well-known examples that fitness to support desirable conduct on the part of citizens, or briefly, to support moral behavior, has served through the ages as a reason for the acceptance of a theory.'

The operation of extra-empirical criteria of theory-evaluation in science is brought to the attention of Margenau and Frank despite rather than by their analytical concerns, which centre upon the empirical virtues of theories. These concerns ensure that extra-empirical criteria are offered a marginal and secondary place in positivist treatments. A more forthcoming reception is afforded by those philosophical schools arising in the early 1960s which struck towards science a historicist rather than analytical stance. J. Agassi was one of the first to address metaphysical criteria of theory-evaluation under this new light. His treatment of the metaphysical roots of scientific problems showed that metaphysics had in history regularly been used to appraise scientific theories. For instance, the continuum theory (the study of the properties of matter, especially elasticity, on the metaphysical assumption that matter is continuous) receives an unfavourable evaluation on the part of anyone whose metaphysical canon embraces atomism.

The problem perceived by Agassi of accounting for the history of science was further and ably addressed by G. Buchdahl. He saw that the historical record of scientific practice offered a test of the completeness or exhaustiveness of the norms composing the methodological canons proposed by philosophers of science. He believed that a canon consisting solely of logico-empirical criteria of theory-evaluation could not pass such a historiographic test:

Although falsifiability, confirmation, and predictive power – all

33 Frank (1957): on the 'scientific' criteria see pp. 348-54; on the 'extrascientific' criteria pp. 354-60.
34 Ibid., p. 359.
36 Agassi (1964), p. 210. That philosophical criteria may be brought to bear upon the evaluation of scientific theories was later argued also by Brush (1974, p. 1,169): 'Those scientists who did suggest that the kinetic theory be abandoned in the later 19th Century did so not because of empirical difficulties but because of a more deep seated purely philosophical objection. For those who believed in a positivist methodology, any theory based on invisible and undetectable atoms was unacceptable.'
involving appeals to observational data, however conceptualized — are important criteria for the choice and acceptability of hypotheses, they are not sufficient, even if one does not go so far as to hold them not to be necessary either.37

Much of Buchdahl’s treatment of this topic is concerned with the development of theories of gravitation, in the conviction that ‘the best defense for any methodological structure like the one here proposed is to observe it at work’;38 he claims to have distilled from the record the criteria upon which theories in this domain were validated:

All such justifications evidently involve an appeal to a supplementary set of ideas: maxims of simplicity and economy; considerations of an esthetic nature; principles of continuity or discontinuity; linkages with general metaphysical notions as for instance ‘the real does not change,’ ‘nothing comes from nothing,’ ‘the effect is equivalent to the cause’; or more generally, maxims like that of homogeneity, affinity (or the ‘analogy of nature’), teleological or alternative preferred explanation schemas, and even certain theological conceptions.39

It is notable that all these treatments regard the category of metaphysical criteria of theory-assessment as fundamental, perhaps on a conceptual par with the category of logico-empirical evaluative criteria, and consign aesthetic criteria — where indeed any is mentioned — to a subclass of the category of metaphysical criteria. The present treatment hopes to show inter alia that an inversion of that hierarchical relation is apt.

---

38 Ibid., p. 213.
39 Ibid., p. 206.
Chapter Three
BEAUTY DISCERNED IN SCIENCE

One of the older research students said the sweetest thing to me after my lecture, that he had never realized there was anything aesthetic in Mathematics till one of my lectures. I was frightfully bucked.1

Sir Nevill Mott

1. Judgements of utility and of aesthetic value

The attainment of predictive power and technological applicability has schooled the observer to regard science as the purveyor of advantage and the vehicle of utility. The capacity of scientific theories which dominates most perceptions is their power to apprehend the causal mechanisms which underlie reality: those who wield that power are enabled to manipulate their surroundings to profit.

This Baconian concern has so overwhelmed all other perspectives on science that attention for any non-utilitarian virtues of theories has been dulled. This is both readily understandable and a matter for regret. The early depiction of science as single-mindedly useful doubtless intended to emphasize and glorify the distinctive capacity of science, distancing it from non-practical or parasitic human pursuits. However, social phenomena which probably include the continued dominance of classicist or literary culture and the increase of leisure which accompanies economic progress turned this utilitarian boast to the discredit of science. The emphasis devoted by partisans of science to its practical virtues were then read as the admission that such were the sole merits to which science could lay claim: views of science emerged in some quarters as the occupation of mechanics, yielding no value higher than the utilitarian.

In these quarters frequent recourse was made to the contrast of science with the arts, seen as the pursuit of contemplation rather than manipulation, as the province of aesthetic rather than utilitarian values. A radical polarization was

1 Mott (1986), p. 36. One gains the impression that this insight has diffused among physicists from generation to generation: Mott recounts that he learned ‘how beautiful physics could be’ from Bohr (ibid., p. 25).
entrenched where a gentle differentiation would have served. The products of science were attributed a monopoly over the utilitarian value denied to art, which for its part was interpreted as the locus of the aesthetic value denied to scientific theories.

The ascription of separate spheres of value to the products of science and of art is the outcome of centuries of conceptualization to which have contributed schools as diverse as positivist philosophy and Romantic literary theory. Anxious to secure recognition for science's benefits to mankind and for the power of scientific reasoning in all domains of the intellect, positivism took to denigrating poetry by comparison. In 1820 the utilitarian T.L. Peacock scorned this genre as a pursuit devoid of practical value in a world constantly bettered by science: 'It cannot claim the slightest share in any one of the comforts and utilities of life in which we have witnessed so many and so rapid advances.' Indeed, the poet 'in the present state of society is [...] a waster of his own time, and a robber of that of others.'

While the denial of utilitarian value to art does not necessarily inhibit the perception of aesthetic value in science, the scientistic dismissal of art as devoid of utility fostered the circumscription of aesthetic value to artistic practice alone. This further step was taken by literary theory. Romanticism commonly attempted to illuminate the nature of poetical works by contrasting them with the dry and unemotional utterances of science: to J. Keats it seemed obvious that matters of fact or science were aesthetically barren. B.R. Haydon recalls hosting a dinner in 1817 for a circle of important Romantic literary figures during which C. Lamb abused the work of Newton as inimical to artistic vision. 'And then he and Keats agreed [Newton] had destroyed all the poetry of the rainbow by reducing it to the prismatic colours. It was impossible to resist him, and we all drank "Newton's health, and confusion to mathematics."' The Romantic view that the scientist's concerns stand in opposition to those of the poet amounted to a denial of aesthetic value to the sciences.

---

2 Brett-Smith and Jones eds (1934), vol. 8, pp. 21-2. Further on the antagonism of positivists towards artistic endeavour see Abrams (1953), pp. 300-3.

3 T. Taylor ed. (1926), vol. I, p. 269. The opinions aired at what Haydon called 'the immortal dinner' - surely one of the most celebrated repasts in the history of literary criticism, attended also by Wordsworth - are considered in Nicolson (1946), pp. 1-5, and Abrams (1953), pp. 303-4.

4 Romanticism's denial of aesthetic value to science did not entail the denial of practical value to the arts. Some Romantic thinkers believed that art had an extrinsic (as opposed to purely aesthetic) value, as a means to moral or social effects beyond itself. Shelley's Defence of Poetry of 1821 was the most elaborately reasoned of all Romantic statements of the moral value of poetry. On this current in Romantic thought see e.g. Abrams (1953), pp. 326-35.
CHAPTER THREE: BEAUTY DISCERNED IN SCIENCE

The tendency to attribute utilitarian value exclusively to the sciences and aesthetic value to only the arts has affected not only the conceptual image of the two domains but also their practice; it has in other terms coloured not only the way in which communities regard the two spheres but also the precepts by which their practitioners pursue them. In particular it has been suggested by some that the products of the two endeavours should be judged by reference to quite separate evaluative categories. On this view the sole category by which the products of science may be assessed is utilitarian, invoking the notion of observational success and a class of criteria similar to the logico-empirical canon codified in the previous chapter. Conversely it is thought that aesthetic categories of appraisal are suitably applied only to artistic creations as contemplative works detached from considerations of practical advantage.

The partition of evaluative categories between the arts and the sciences along these lines has dominated most treatments of science and many broader discussions of culture in the last two hundred years. The present treatment believes nonetheless that the partition is inappropriate, and depends to some measure upon an excessively simple view of scientific practice and to a further extent upon a narrow definition of the class of objects deemed susceptible of aesthetic perception. The immediate concern of the present treatment is to instil doubts in the partition of evaluative categories between the arts and the sciences, and in particular to suggest that the products of science as well as those of art are susceptible to aesthetic evaluation.

This suggestion will be entrenched in several stages in the course of this chapter and the next. First the roots will be retraced of the notion of intellectual beauty in philosophical aesthetics; then some evidence will be adduced to support the contention that scientific and other intellectual communities discern beauty in the products of research. Finally, the next chapter will explore the philosophical considerations underpinning the application of aesthetic evaluative categories to scientific theories.

One of the earliest sophisticated modern proponents of the notion that intellectual constructs may possess aesthetic value was F. Hutcheson, a member of the eighteenth-century British empiricist school. His writings on this topic are of especial interest to the present treatment, for the following reason. The next chapter will retrace partly to Hutcheson the view of aesthetics as the study of disinterested perception, an interpretation upon which the present study of the

---

5 This suggestion is advanced by e.g. Feigl (1970), pp. 9-10, whose views will be considered in chapter 5 below.
perception of beauty in scientific theories will largely be based. It will be suggested also in the next chapter that, as occurred in Hutcheson's work, the interpretation of aesthetics as the study of the disinterested mode of attention rather than as a differently constituted investigation leads naturally to a broadening of the class of entities designated as objects of aesthetic perception: in particular this class is widened to include not only bona fide works of art but also non-artistic objects and indeed intellectual constructs like scientific theories. Appeal to the interpretation of aesthetics as the study of disinterested perception will therefore facilitate the present treatment's study of aesthetic criteria of theory-evaluation in science in two ways: first, as explored in the next chapter, by providing the philosophical stance from which to consider the perception of aesthetic virtues in scientific theories; second, as outlined in the remainder of this chapter, by tabling the presumption that aesthetic evaluative categories should be extended to the study of intellectual constructs like scientific theories. The *ordo docendi* chosen here for the exposition of the implications of the view of aesthetics as the study of disinterested perception departs from the *ordo inveniendi* in which these implications would be analytically encountered: hence this chapter will consider the remoter implications of this view - the extension of aesthetic categories to intellectual constructs - while the next chapter will consider the central doctrines of this view from which descends *inter alia* the broadening of the category of objects of aesthetic perception to embrace the class of intellectual constructs.

2. Hutcheson's discussion of intellectual beauty

Hutcheson's account of beauty in intellectual constructs derives unbrokenly from his more general aesthetic theory: the objective qualities to which he attributes the cause of the occurrence of ideas of beauty in the observer are in his theory the same independently of whether the object of perception is an intellectual construct or a material object, such as a natural scene or a work of art.6

Hutcheson's theory of beauty stands in the context of representative

---

6 Few critical discussions of Hutcheson's aesthetics have yet appeared, but the edition of Hutcheson (1725) utilized in this section includes an introduction, notes and bibliography by Kivy.
realism, of which the central presupposition was that the qualities of objects are distinct from, and in fact the causes of, 'ideas', the only immediate materials of sensory awareness. Accordingly for Hutcheson beauty is an idea occasioned in the mind by a quality of external objects: he states that 'the word beauty is taken for the idea raised in us, and a sense of beauty for our power of receiving this idea.' In further characterizing the nature of the idea of beauty Hutcheson appeals to the distinction between ideas of 'primary' and ideas of 'secondary' qualities, drawn by J. Locke as follows:

The ideas of primary qualities of bodies are resemblances of them, and their patterns do really exist in the bodies themselves; but the ideas produced in us by these secondary qualities have no resemblance of them at all. There is nothing like our ideas, existing in the bodies themselves. They are in the bodies, we denominate from them, only a power to produce those sensations in us: and what is sweet, blue, or warm in idea, is but the certain bulk, figure, and motion of the insensible parts in the bodies themselves, which we call so.\(^7\)

Whereas the primary qualities are inseparable from matter and are found in every part of it, the secondary qualities are not intrinsic qualities but merely powers in the objects to produce sensory effects in us by means of the primary qualities. In a passage clearly inspired from the Lockean formulation, Hutcheson sets out a view of beauty as a secondary quality consisting of the reaction of the observer to certain of the object's primary qualities:

Let it be observed that by absolute or original beauty is not understood any quality supposed to be in the object which should of itself be beautiful, without relation to any mind which perceives it. For beauty, like other names of sensible ideas, properly denotes the perception of some mind; so cold, hot, sweet, bitter, denote the sensations in our minds, to which perhaps there is no resemblance in the objects which excite these ideas in us, however we generally imagine otherwise.\(^8\)

Hutcheson here construes beauty as the response of the aesthetic perception to objective characteristics of the perceived.

Having specified what kind of entity are ideas of beauty, Hutcheson turns to investigate the quality of objects which causes their occurrence in the human mind. 'Since it is certain', he writes, 'that we have ideas of beauty and harmony, let us examine what quality in objects excites these ideas, or is the occasion of

---

\(^7\) Hutcheson (1725), p. 34; emphasis in the original.

\(^8\) Locke (1690), p. 137. I have simplified the typography of this passage.

\(^9\) Hutcheson (1725), pp. 38-9; emphasis in the original.
Hutcheson’s examination quickly yields the required property of objects of perception:

The figures which excite in us the ideas of beauty seem to be those in which there is uniformity amidst variety. [...] What we call beautiful in objects, to speak in the mathematical style, seems to be in compound ratio of uniformity and variety; so that where the uniformity of bodies is equal, the beauty is as the variety; and where the variety is equal, the beauty is as the uniformity.11

In Hutcheson’s view the property of ‘uniformity amidst variety’ causes those ideas of beauty which are formed in the mind through the perception of scenes in nature or works of art.

A considerable part of the remainder of Hutcheson’s treatise on aesthetics is concerned to provide examples of the application of this explanatory principle to the perception of various classes of objects. Hutcheson clearly hopes that those objects which in his view can be shown to possess ‘uniformity amidst variety’ to the greatest degree will also be those which the reader naturally finds most beautiful. He considers as a preliminary example the set of the regular polygons, perhaps because the number of their sides permits a clearly comprehensible and easily quantifiable measure of the degree to which they exhibit uniformity amidst variety, and thus a measure of the potency with which their perception will cause the idea of beauty to be conceived in the observer:

First, the variety increases the beauty in equal uniformity. The beauty of an equilateral triangle is less than that of the square, which is less than that of a pentagon, and this again is surpassed by the hexagon. [...] The greater uniformity increases the beauty amidst equal variety in these instances: an equilateral triangle, or even an isosceles, surpasses the scalenum; a square surpasses the rhombus or lozenge.12

The modern reader finds the project of ranking geometrical figures according to a degree of beauty which varies proportionally with the number of their sides artificalicious and gratuitous: artificalicious because there is severe doubt whether the notion of beauty can be linked in such a pedestrian manner to a numerical variable, and gratuitous because it is no longer considered the task of an aesthetic theory to order geometrical figures by degree of beauty.

Whilst Hutcheson fails to impart lasting value to his treatise in those passages which speculate on quantitative measures of beauty, his remarks on
intellectual beauty attain greater significance. Hutcheson’s anxiety to extend the scope of aesthetic theory to embrace the aesthetic dimensions of intellectual constructs is demonstrated when his treatment turns to consider the ways in which the property of ‘uniformity amidst variety’ may be said to be encountered in the cultivation of the empirical sciences. His attention is devoted mainly to astronomy, where he finds his views most clearly illustrated.

Hutcheson draws a distinction which is obviously fundamental to any analysis of the aesthetic appeal of science: he differentiates between the beauty of the systems which constitute the subject-matter of science and the beauty intrinsic to the theories which model or study those systems. Where however a less subtle investigator setting out to study the circumscribed topic of the beauty of theories would simply have dismissed the first form of beauty as irrelevant to the project, Hutcheson sees there too material useful to the consideration of the aesthetic aspects of scientific theories. The progressive unfolding of Hutcheson’s treatment of the matter suggests that he believes the aesthetic perception of the natural systems studied by scientific theories to be more complex than the process of merely observing a landscape and drawing from it a sensation of beauty.

According to Hutcheson the idea of beauty produced by contemplation of the natural systems studied by science is not reducible to the similar idea obtained from the unmediated observation of a natural scene: the former experience is crucially mediated by the scientific theory which alone allows us to grasp the structure of the system and to recognize beauty in it. The sensation of beauty which one feels under those circumstances could not be attained without recourse in the very act of perception to the theory which models the natural systems constituting the object of perception. The aesthetic perception of a natural system under those circumstances is a perception of the system through the mediation of theory. One might term it a theory-laden aesthetic perception of a natural phenomenon by similarity with the theory-laden observation of so-called empirical qualities of which the philosophy of science has long debated: in each case the quality attributed to the object is perceived not unmediatedly but through the crucial interpretative medium of relevant theory. Hutcheson thus appears to suggest that the theory-laden aesthetic perception of a natural system which constitutes the subject-matter of a complex of scientific theories is essentially different from the common-sense aesthetic perception of a natural scene to which our access is pre-scientific.

Hutcheson thus envisages three forms of aesthetic perception, each of which is further removed from nature and more heavily dependant upon
conceptualization than is the previous one, but all of which are apt to be occasioned by the scrutiny of the same natural object: the unmediated aesthetic perception of natural scenes, the aesthetic perception of natural systems as these are comprehended by the mediating scientific theory, and the aesthetic perception of scientific theories themselves. These three increasingly abstract forms of aesthetic perception find their correspondent in the three stages of Hutcheson's treatment of the astronomical example.

Hutcheson first acknowledges that celestial bodies show in the regularity of their shapes and their orbits great aesthetic appeal:

The forms of all the great bodies in the universe are nearly spherical, the orbits of their revolutions generally elliptic, and without great eccentricity, in those which continually occur to our observation. Now these are figures of great uniformity, and therefore pleasing to us.13

The perception of the sphericity of the celestial bodies or the near-circularity of their orbits may well be attained by observation unmediated by theory, though perhaps not by the efforts of an earth-bound observer alone. Hutcheson moves to the second level of aesthetic perception in finding the origin of beauty not only in the shapes and motions of celestial bodies but also in the understanding of celestial configurations in terms characteristic of the astronomical theory which the observer employs to interpret the natural system before his or her eyes. This new sensation of beauty is yielded by theory-mediated perception and hence is not reducible to the sensation experienced by merely gazing at the night sky. In Hutcheson's example this sensation is gained from contemplation of the constancy or the interrelations of the quantitative parameters characteristic to the theory's structure, which are not offered to the perception save through the suggestion of theory. Hutcheson refers to terms such as masses, distances and periods of revolution, none of which is an object of unmediated perception but each of which is apprehended through understanding of the theory. The allusion in the following passage to the routine calculations of the astronomer suggests that Hutcheson envisaged perception of this form of beauty to be characteristic to the expert practitioner, who alone is able to apprehend the necessary mathematical variables and relations:

Further, to pass by the less obvious uniformity in the proportion of their quantities of matter, distances, times of revolving, to each other, what can exhibit a greater instance of uniformity amidst variety than the constant tenour of revolutions in nearly equal times,

---

13 Ibid., p. 42.
in which each planet around its axis, and the central fire, or sun, through all the ages of which we have any records, and in nearly the same orbit? [...] These are the beauties which charm the astronomer, and make his tedious calculations pleasant.  

The third stage of Hutcheson's treatment turns not to the sense of beauty afforded by the contemplation of nature, nor yet the sense gained from the contemplation of nature mediated by the scientific understanding, but the sense obtained from the perception of the products of science, or scientific theories. Hutcheson clearly believes that some theories possess the property of 'uniformity amidst variety' to exemplary degree. It is in theories of great generality which are known with certainty, or 'universal truths demonstrated', that Hutcheson discerns the greatest capacity for aesthetic appeal: there is no kind of beauty 'in which we shall see such an amazing variety with uniformity, and hence arises a very great pleasure distinct from prospects of any further advantage.' The reason for this is that in such theories 'we may find included, with the most exact agreement, an infinite multitude of particular truths, nay, often a multitude of infinities.' He contrasts the strength of the idea of beauty produced in those who contemplate such a theory with that afforded by the perception of a generalization which is merely inductive and thus not known with certainty:

Let us compare our satisfaction in such discoveries with the uneasy state of mind when we can only measure lines, or surfaces, by a scale, or are making experiments which we can reduce to no general canon, but are only heaping up a multitude of particular incoherent observations. Now each of these trials discovers a new truth, but with no pleasure or beauty, notwithstanding the variety, till we can discover some unity or reduce them to some general canon.

Hutcheson states that the beauty which derives from the contemplation of generalizations proved with certainty is to be found notably in geometry; he does not however omit to mention an example from the empirical sciences of a theory capable of producing in the observer an equally strong idea of beauty:

There is another beauty in propositions when one theorem contains a great multitude of corollaries easily deducible from it. Thus there are some leading or fundamental properties upon which a long series of theorems can be naturally built. Such a theorem is the 35th

---

14 Ibid., pp. 42-3.
15 Ibid., p. 48.
16 Ibid. The idea that the contemplation of such constructs may bring aesthetic gratification which is 'distinct from prospects of any further advantage' will play an important part in the definition of the attitude of disinterested attention which the next chapter will present as characteristic of one construal of aesthetics.
17 Ibid.
18 Ibid., p. 49.
of the 1st Book of Euclid, from which the whole art of measuring right-lined areas is deduced by resolution into triangles which are the halves of so many parallelograms [...]. In the search of nature there is the like beauty in the knowledge of some great principles or universal forces from which innumerable effects do flow. Such is gravitation in Sir Isaac Newton’s scheme.19

Maintaining consistency with the principle which sees as the cause of ideas of beauty the property of ‘uniformity amidst variety’, the beauty which Hutcheson discerns in theories of empirical science springs from their generality and unifying power. As the following chapters will illustrate, this feature of theories has been cited by practising scientists from the time of Hutcheson to the present as a central component of their notions of theoretical beauty: Hutcheson’s theory of intellectual beauty to this extent captured elements of the aesthetic sensibility and predispositions of the working scientist.

Hutcheson’s treatment of the beauty of scientific theories serves as an example of an intellectual itinerary from the most general principles of an aesthetic theory – such as the principle that the cause of the idea of beauty in the mind of the observer is the object’s possession of uniformity amidst variety – to the formulation of the problems attendant upon the notion that scientific theories may display and be evaluated on their possession of aesthetic value. While the details of Hutcheson’s approach are not shared by the present treatment, and hence in this context a further examination of Hutcheson’s views would be of largely historiographic interest, the present section has aimed at two results: illustrating the way in which the presumption that scientific theories possess aesthetic value was reached by a representative exponent of an influential aesthetic school, and setting the conceptual scene for what is to ensue. In particular, the next chapter will penetrate to a level of analysis deeper than that attained in the above exposition of Hutcheson’s views, and investigate the nature of the aesthetic outlook which permitted and fostered Hutcheson’s speculations on the beauty of intellectual constructs. The conviction that one may or ought to search for beauty in intellectual constructs and more especially in scientific theories will there be linked to the aesthetic outlook characterized by the notion of disinterested attention. Before that issue is broached, the remainder of this chapter will consider some other expressions of the view that intellectual constructs display aesthetic value.

19 Ibid., pp. 50-1; emphasis in the original.
3. Some twentieth-century invocations of beauty in science

Today the application of aesthetic categories to intellectual constructs is considerably less familiar to aesthetic theory than it was at the time of Hutcheson. H. Osborne laments this decline, which he attributes to changes in the concerns of the philosophical community rather than to any lack of value of the notion:

Nowadays the concept of intellectual beauty is not, I believe, commonly repudiated as much as neglected; few of the standard works on aesthetics pay more than lip-service to it and I know of none which has either attempted a deep analysis or given to it equal weight with sensory beauties in the framing of general aesthetic concepts.²⁰

If twentieth-century writers in systematic aesthetics have tended to neglect the notion of intellectual beauty, greater interest in this topic has been demonstrated by scientists. Later chapters will illustrate and study scientists' use of aesthetic features in their practice of theory-assessment; but even where scientists have stopped short of suggesting explicitly that such features be used to evaluative ends, they have frequently voiced the conviction that scientific theories exhibit aesthetic features which are an occasion of delight.

For instance, H. Poincaré attributed great importance to the presence of certain aesthetic features in theories of his domain of physics, to the extent of appearing to believe that the distinguishing attribute of the mathematical mind is not logical but aesthetic.²¹ Poincaré's explanation of the distinction between the intellectual beauty encountered in the mathematical investigation of nature and the 'sensory' beauty experienced in merely regarding natural scenes recalls the corresponding distinction traced by Hutcheson:

If nature were not beautiful, it would not be worth knowing [...]. Of course I do not here speak of that beauty which strikes the senses [...] I mean that profounder beauty which comes from the harmonious order of the parts and which a pure intelligence can grasp. [...] Intellectual beauty is sufficient unto itself, and it is for its sake, more perhaps than for the future good of humanity, that the scientist devotes himself to long and difficult labours.²²

Elaborating in another work upon the mechanism by which the 'profounder

---

²⁰ Osborne (1964), p. 160. This essay somewhat ruefully observes that one of the few practices which today invokes the notion of intellectual beauty is the study of chess. Osborne's more specific remarks on beauty in physical theories are contained in his (1984).


beauty’ can be grasped by a ‘pure intelligence’, Poincaré suggests that an aesthetic sensibility is exercised in scientific practice:

It may appear surprising that sensibility should be introduced in connexion with mathematical demonstrations, which, it would seem, can only interest the intellect. But not if we bear in mind the feeling of mathematical beauty, of the harmony of numbers and forms and of geometric elegance. It is a real aesthetic feeling that all true mathematicians recognize, and this is truly sensibility.\(^2\)

Like Hutcheson in quite another historical context, Poincaré moves from the acknowledgement that a sense of beauty is attained in the contemplation of certain intellectual constructs to the investigation of the properties of those constructs which are apt to produce the sense in the mind of the observer. He reaches a conclusion strikingly similar to that of Hutcheson:

What are the mathematical entities to which we attribute this character of beauty and elegance, which are capable of developing in us a kind of aesthetic emotion? Those whose elements are harmoniously arranged so that the mind can, without effort, take in the whole without neglecting the details. This harmony is at once a satisfaction to our aesthetic requirements, and an assistance to the mind which it supports and guides.\(^3\)

Poincaré’s attribution of the idea of beauty to the perception in an intellectual construct of a harmonious arrangement which enables the mind to apprehend the whole and equally the details within the whole recalls Hutcheson’s identification of the cause of ideas of beauty with the object’s possession of ‘uniformity amidst variety’.

Poincaré’s close contemporary P. Duhem placed no less stress on the beauty of physical theory, in a passage tinged by somewhat Whiggish historiography:

It is impossible to follow the march of one of the great theories of physics, to see it unroll majestically its regular deductions starting from initial hypotheses, to see its consequences represent a multitude of experimental laws, down to the smallest detail, without being charmed by the beauty of such a construction, without feeling keenly that such a creation of the human mind is truly a work of art.\(^4\)

A few years later the same portrayal of scientific investigation as an endeavour rich in aesthetic or artistic tones, which contribute to the practitioner’s conviction that science is a valuable and noble product of the human mind, was pursued by

\(^{23}\) Poincaré (1908), p. 59.
\(^{24}\) Ibid.
B. Russell:

Mathematics, rightly viewed, possesses not only truth, but supreme beauty - a beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or sculpture, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show.  

While the present treatment advances no suggestion that such brief statements bear the authority necessary to underpin the model of scientific rationality to be outlined in the chapters to follow, by such testimonials the suspicion gains purchase in the mind of the philosopher that the aesthetic appreciation and evaluation of scientific theories is a genuine component of actual scientific practice. The degree to which this component may be said to observe the methodological injunctions expressed by canons of scientific rationality remains to be determined; but the existence of a subject-matter for the succeeding investigation is thereby assured.

4. The literary reception of scientific theories

Aesthetic value is discerned in scientific theories not only by scientists but also by literary and artistic practitioners. The arts have throughout their history devoted attention to the findings and theories of the sciences, and have frequently identified in scientific theories elements endowed of aesthetic value which they have sought to incorporate into their own creations. The aesthetic estimation of elements of scientific doctrines within literary or artistic works suggests that scientific theories may possess aesthetic value: to this extent the findings of literary criticism conform with the pronouncements of scientists themselves in supporting the contention that scientific theories are susceptible of aesthetic evaluation.

One of the earliest literary critics to address in systematic fashion the task of analyzing the references to scientific beliefs in works of literature was I.A.

26 B. Russell (1917), p. 49.
27 See Richards (1926).
Richards. Since his time studies of this topic have greatly proliferated. The amount of data and analysis therein accumulated renders it unnecessary for this section to provide more than brief examples of the aesthetic estimation of scientific beliefs in literary works.

That literature has frequently drawn its aesthetic devices from current or recent natural philosophical or scientific beliefs is indubitable. For instance, the imagery of the centre and the circumference of the circle are persistent aesthetic devices in seventeenth-century English poetry and especially in the works of J. Donne, J. Milton and A. Marvell. These devices are often explicitly drawn from the Aristotelian and Ptolemaic cosmology which had prevailed earlier in the century and which attributed to the orbits of celestial bodies the circular form considered the sole figure appropriate to the perfection of the superlunary region. Milton transposes the circle from planetary astronomy to Christian poetry in describing the angels' celebration of God's pronouncement 'This day I have begot whom I declare My only Son':

That day, as other solemn days, they spent
In song and dance about the sacred hill,
Mystical dance, which yonder starry sphere
Of planets and of fixed in all her wheels
Resembles nearest, mazes intricate,
Eccentric, interwoven, yet regular
Then most, when most irregular they seem,
And in their motions harmony divine
So smooths her charming tones, that God's own ear
Listens delighted.

It is clear that in this passage Milton attributes to the angels in the celebratory dance motions drawn from the domain of cosmology: the circular motions attributed by Ptolemaic astronomy to celestial bodies are transposed from figures of astronomical theory to poetical trope, and as such are beheld as vehicles of aesthetic value. Milton's aesthetic delight is stimulated by the perception not of the motions of the angels in the imagined dance, nor yet by the motions of the

---

27 The critical discussion of literary references to scientific theories has been surveyed by Schatzberg, Johnson and Waite eds (1987), which has largely superceded Dudley ed. (1968). Although such discussions seem to have been pursued most keenly in English-language literary criticism, the same topic has attracted attention in other cultures; the Italian debate on literary references to science, for instance, is charted by Battistini ed. (1977), esp. pp. 222-34.

28 Further on the incorporation of cosmological imagery in seventeenth-century English poetry see Nicolson (1950), pp. 47-80, and (1956), pp. 1-79.

bodies in the heavens, but by the perception of the astronomical theory which attributes such motions to the celestial bodies. That the object of Milton's aesthetic delight is the theory rather than any physical manifestations of movement is suggested by his reference not to the motions of the celestial bodies as they appear in the sky to the observer, which are 'intricate, eccentric, intervolved' and thus scarcely regular or harmonious, but to their actual motions which are 'regular then most when most irregular they seem.' Any regularity of the celestial motions to which Milton refers is invisible to the observer of the physical phenomena and can be apprehended only by turning one's gaze from the heavens to the domain of physical theory: it is in the physical theory rather than in the night sky that any regularity resides. The harmony in which Milton delights is proper not to the observed motions of the celestial bodies but to the structure of Ptolemaic theory. It is hence the astronomical theory itself, or the world view which is underpinned by that theory, which is in Milton's eyes the source of aesthetic value.

If Milton's poetry expresses the serenity and confidence imbued by the aesthetic delight at the perception of a secure world-view, the poetry of Donne can express aesthetic horror at the overthrow of that cosmology. Donne voices the apprehension of his contemporaries at the realization that natural philosophy no longer constitutes the model of harmony and propriety:

And new Philosophy cal's all in doubt,
The Element of fire is quite put out;
The Sunne is lost, and th'earth, and no mans wit
Can well direct him, where to looke for it.
And freely men confesse, that this world's spent,
When in the Planets, and the Firmament
They seeke so many new; they see that this
Is crumbled out againe to his Atomis.
'Tis all in pieces, all cohaerence gone;
All iust supply, and all Relation.31

Again, the object of the poet's perception, and the source of Donne's horror as much as of Milton's satisfaction, is not the physical appearance of the universe but rather the philosophical theories which model that appearance. The stimulus responsible for a community's transition from the state of satisfaction expressed by Milton to the state of anxiety voiced by Donne could not have been a new observation of the behaviour of the universe: clearly when Donne laments that this world 'is crumbled out againe to his Atomis' he refers not to any newly-

31 The First Anniuersary: An Anatomy of The World, ll. 205-14 (Manley ed. 1963, p. 73). This poem was written in 1611. Further on Donne's cosmological imagery see Nicolson (1956), pp. 58-79.
performed observations of the universe but rather to the perception of new theories of the universe. The aesthetic object of Donne’s poem is thus a conceptual construct, a natural philosophical theory of which he was appraised and which he saw as an object of aesthetic revulsion.

While English poets of the seventeenth century displayed aesthetic responses to the structure of Aristotelian and Ptolemaic cosmology and to its overthrow, the aesthetic attention of those of the succeeding century was occupied partly by the world-view erected by Newtonian physics. Both the deterministic mechanics of the *Principia* and the theory of light of the *Opticks* attracted literary responses: the conception of the universe as a clockwork mechanism and of white light as a mixture of the spectral colours provided basic imagery and symbolism for much eighteenth-century poetry. To the extent to which they were utilized as poetic figures, these conceptions were perceived as bearing aesthetic value.

Scientific theories have continued to attract aesthetic responses to this day. Darwin’s theory of evolution was perceived as a construct of rich aesthetic value. In the twentieth century the theory of relativity has provoked a wealth of literary and artistic discussion, which has focused yet again not on the phenomena postulated by the theory but on the theory’s assertions and the rupture which it wrought on general culture.

Literary criticism finds that scientific theories constitute an object of aesthetic perception in the literature of all ages. If theories are deemed appropriate objects of aesthetic perception by literary practitioners, the suggestion appears more plausible that scientific communities too should record and react to the aesthetic features of theories which they encounter. The principles of philosophical aesthetics under which such a perception of scientific theories is possible form the object of the next chapter.

---

33 The literary responses to Darwinism have been surveyed by Beer (1983), pp. 104-258, and Levine (1988).
34 The reaction of the literary-artistic community to the theories of relativity is extensively described by Friedman and Donley (1985), esp. pp. 67-109; further on the relation between Einstein’s work and the humanities see e.g. Ryan ed. (1987).
5. The perception of beauty in abstract objects

While the interest of the present treatment lies in scientific methodology rather than cognitive psychology, it is necessary to adduce some explanation of how it may be possible for the observer to discern aesthetic value in an abstract object such as a scientific theory. Clearly it is not adequate to suggest that such intellectual beauty is perceived directly through the exercise of the normal or external senses, as the beauty of a painting would presumably be perceived.

Hutcheson suggested a solution by positing the existence alongside the 'external senses' – which are held responsible for the perception of ideas of beauty in external bodies – of an 'internal sense' apt to conceive ideas of beauty from the contemplation of abstract objects.35 He cites evidence from musical aesthetics: 'In music we seem universally to acknowledge something like a distinct sense from the external one of hearing, and call it a good ear.'36

It is however not necessary to postulate an extra human sense for the perception of beauty in abstract objects. It is sufficient to suppose that the observer perceives elements of beauty not in the abstract object itself but in a token or representation of the object to which the aesthetic features of the object are transposed. For instance, a musical composition is universally deemed to possess aesthetic features, but it is at least plausible to suggest that those who contemplate such features perceive them not in the composition in itself, which is an abstract entity, but in some performance or other representation of the composition. Musicology implicitly acknowledges an operation similar to this transposition when it refers to the aesthetic features of a piece upon perception of the aesthetic features of one of its performances.

Similarly it is plausible to suggest that the aesthetic features of scientific theories are perceived not directly in the abstract entity constituted by the theory but in some formulation of the theory on paper. By this device is averted the necessity of postulating the existence of human sensory modes outside those externally activated.

35 Hutcheson (1725), p. 34.
36 Ibid., p. 35; emphasis in the original.
Chapter Four
THE AESTHETIC STANCE

1. Two views of aesthetics

The previous chapter outlined some expressions of the view that aesthetic value is to be found in intellectual constructs and especially in scientific theories. These expressions were used to suggest that there exists in intellectual communities — both within scientific circles and more widely — a practice of discerning beauty in science as well as in other objects of the intellect. Since the purpose of the previous chapter was merely to advance this suggestion in the mind of the onlooker, and thereby to gain legitimacy for the study of aesthetic features of science, little reference to philosophical aesthetics was required. The present chapter on the contrary will turn to considerations taken from aesthetic theory.

The purpose of this chapter is to delineate a view of aesthetics which exhibits two characteristics: it possesses a distinguished history in aesthetic speculation, and it offers an outlook particularly well suited to the investigation of aesthetic features of scientific theories. This chapter will therefore establish the foundations in philosophical aesthetics for the study of aesthetic criteria of theory-assessment which constitutes the principal object of this work.

The discipline of aesthetics may broadly speaking be defined in two ways. It may be defined as the study of some of the features of a specified class of objects of perception, or as the study of a specified mode of attention or perception in the beholder. Definitions of the former type may be termed objective in as far as they stipulate the class of objects deemed susceptible of aesthetic perception, whilst definitions of the second type are subjective in the sense that they codify the mode of perception which is constitutive of aesthetic perception. Use here of the terms 'subjective' and 'objective' is not intended to suggest any difference between the degrees to which the two definitions admit of
Definitions of the subjective type possess advantages over objective definitions for both the philosophical fertility of aesthetics in general and the study of the aesthetic features of scientific theories in particular. Consider first the approach of aesthetic theorists who propose to define their discipline by reference to a class of objects. They will tend to identify in their mind a nucleus of that class composed of what they take pre-paradigmatically to be 'aesthetic objects', and define the class as the set of objects which agglomerate around that nucleus. This procedure exhibits two possible shortcomings for present purposes.

First, definitions of aesthetics of the objective type tend to betray undue contamination by art criticism in basing their stipulation of the class of aesthetic objects upon the notion of the work of art, which is seen as the quintessential 'aesthetic object'. Discussions in aesthetics where the discipline is held to be objectively defined consequently tend to decay into disquisitions on what should count as a work of art and hence as an aesthetic object, during which one commonly witnesses the trading among the interlocutors of features deemed indicative of 'good art'. While the comparative evaluation of the aesthetic effectiveness of works of art is an enterprise of interest for the discipline of art criticism, it ought presumably to be pursued downstream of the foundation of philosophical aesthetics rather than at the outset of aesthetic discussion.

The attribution of excessive weight to the work of art in the definition of the class of aesthetic objects is perhaps a particular instance of the second possible shortcoming of the procedure by which the discipline of aesthetics is objectively defined. This concerns the speed of response of aesthetic discussion to conceptual innovation. Definitions of a practice or phenomenon which are based on a preparadigmatic intuition of the nature of the object are almost invariably affected by the problem of conservatism. Frequently the term 'intuition' to which appeal is

---

1 The history of subjective and objective theories of aesthetics is presented in e.g. Listowel (1967), pp. 3-156.

2 This approach to the definition of aesthetics is pursued by Beardsley (1961). That work first signals its endorsement of objectivist definitions of the discipline: 'It seems to me useful for aesthetics to have a generic term to mark out, though vaguely, the objects within its field of interest [...] say the neologism "aesthetic object"' (ibid., p. 177). Beardsley then associates the notion of aesthetic object indissolubly to the work of art: 'I stipulate that all musical compositions are aesthetic objects, all literary works are aesthetic objects, all works of plastic art are aesthetic objects,...etc.' (ibid.). He lastly attempts to characterize what makes a work of art good (ibid., pp. 184-5), a notion around which the class of aesthetic objects will coalesce. It is unlikely that Beardsley's approach would have been able gracefully to resolve the polemic raised over the Tate Gallery's purchase of the controversial abstract work to which allusion will be made later in this section: would that object have counted as 'a work of plastic art' or not?
made in such a procedure is a euphemism for common belief engendered by prolonged exposure or persuasion: the supposedly unproblematic intuition thus represents no more than a prejudice entrenched by past speculation, and offers an insecure basis for a philosophically fertile approach to the definiendum. This danger potentially besets the identification of a nucleus of the class of ‘aesthetic objects’: a theorist’s view of what counts as a quintessential aesthetic object may result from an upbringing which attributed instinctive preference to a particular past aesthetic theory. If so, this procedure imposes a time-lag between advances in aesthetic speculation and definitions of the subject-matter of aesthetics: the definition of aesthetics accepted at some time would be determined by the beliefs of a former time and hence lag behind current aesthetic thought. In particular, were daring speculation to suggest that a new set of objects not hitherto included in the definitorial class be treated by aesthetic theory, the discipline defined by reference to that class would experience difficulty in adjusting to the innovation.

The approach to aesthetic theory predicated upon objective definitions is in these two ways potentially deleterious to the philosophical fertility of aesthetic theory in general because it imposes a pre-emptive delimitation of its subject-matter, circumscribing the breadth of its analysis even before the analysis has been undertaken. Definitions of aesthetics of the subjective type do not present this danger, since they characterize the discipline as the study not of features of a certain class of objects, but rather of features of a certain and distinctive mode of attention which may be directed upon objects of perception of any specification. Subjective definitions are thus less prone than are their objective counterparts to prejudice the question of the range of objects to which aesthetic categories may be applied: subjective definitions of aesthetics specify the cast of mind with which the aesthetic observer turns to his or her object, but approach with an open mind the issue of what the nature of this object may be.

In view of the fact that aesthetic concerns are closer to the public’s heart than are those of other branches of philosophy such as metaphysics or epistemology, it is perhaps not surprising that the contention between objective and subjective definitions of aesthetics has at times erupted into popular debate. Some of the discussions have manifested the shortcomings of objective definitions

---

3 The danger that the definition of a notion by reference to a pre-paradigmatic intuition may lead to an excessively conservative approach to the study of that notion is apparent in disciplines outside aesthetics, and specifically in the philosophy of science. J.R. Brown (1980, pp. 238-40) criticizes as an approach subject to this danger the reliance of Laudan (1977, pp. 160-1) on pre-analytic intuitions in the characterization of scientific rationality.
of aesthetics to which reference was made above. In 1972 the Tate Gallery of London acquired *Equivalent VIII*, a composition by the contemporary U.S. minimal sculptor C. Andre consisting of 120 identical ready-made commercial firebricks arranged in a low parallelepiped. Four years after its purchase, this work became the object of a controversy raised by the British press and reported world-wide. Critics alleged that an arrangement of bricks did not constitute a suitable object of expenditure for an art gallery; defenders of the acquisition maintained that such criticism presupposed an excessively circumscribed definition of what counts as a work of art, and suggested that a structure of everyday objects like bricks could carry aesthetic value. The substance of the controversy was tersely captured in *The New York Times*’s headline ‘Tate Gallery Buys Pile of Bricks – Or Is It Art?’.

At the centre of this controversy lay the two approaches to the definition of aesthetics outlined above. Prescribing an objectivist definition of aesthetics, critics of the Tate’s purchase would doubtless have insisted that the gallery’s resources be devoted to paradigmatic art-works like paintings or sculptures, presumably because these pieces fall within the class of objects which the critics deemed susceptible to aesthetic perception. Defenders of the gallery’s acquisition policy argued on the contrary that any artefact could be an object of aesthetic perception, and that the arrangement of bricks was in principle as worthwhile an object of aesthetic contemplation as a Constable. Underlying this broad-mindedness is a notion of aesthetics as the study of the properties of perception rather than of the properties of a certain class of objects of perception.

It seems a particularly barren philosophical project to attempt to determine whether a given object is to count as a pile of bricks or a work of art. One may seek to evade this morass by refusing to constitute the discipline of aesthetics in terms of the class of objects of perception to which aesthetic categories will be applied, and to choose instead to define the discipline on the basis of the attitude of the subject of perception.

If definitions of aesthetics of the objective type possess certain shortcomings for the pursuit of aesthetic speculation in general, they exercise an especially severe restrictive effect upon the study of the aesthetic features of scientific theories. Objective definitions strive to demarcate the boundaries of the class of aesthetic objects in the light of extant ideas of what constitutes paradigmatic objects of aesthetic perception and of preceding speculation in aesthetic theory. As

---

4 For details of Andre’s work and the beginnings of the controversy see Alley (1981), pp. 11-2.

5 The article of which this was the headline is Semple (1976).
section 1 of the previous chapter showed, the study of aesthetic features of scientific theories hinges—in opposition to these concerns—upon a negation of the classical boundaries of the aesthetic. The study of the aesthetic features of scientific theories depends upon an unprecedented transcending of the class of objects to which it has customarily been deemed that aesthetic categories may appropriately be applied: this enterprise must guard itself against strangulation at birth by the imposition of an objective definition of aesthetics which unwarrantedly circumscribes the class of the objects of aesthetic perception.

For considerations of this kind, the present treatment will seek to characterize the study of aesthetics by the subjective route. The next section will outline a conception of aesthetics as the study of a certain perceptual mode: the mode characterized by a lack of utilitarian interest in the object of perception. Section 3 of this chapter will retrace the historical origins of this view in the works of eighteenth-century English empiricist theorists.

2. The study of disinterested attention

Here is a theory about the mode of attention which is activated in aesthetic perception. There are perhaps many modes of attention which one may assume in the act of perception. These modes will typically be characterized by the aims or interests by which the perception is animated or in view of which it is conducted. One may for instance gaze upon a gem with the aim to valuate it or regard a chess-board with the interests of Black at heart. According to the theory developed here, neither of these modes of attention is the kind which is activated in aesthetic perception: the mode of attention characteristic of aesthetic perception is disinterested.

‘Disinterestedness’ denotes an attitude of utilitarian neutrality or purposelessness towards an object of perception. As an object of disinterested perception, an object is regarded not as means to some end external to itself but as the self-contained end of a particular enterprise, the enterprise of perception. The disinterested stance towards an object dwells upon not its susceptibility to exploitation in the furtherance of some aim, but its intrinsic structure and significance. Under disinterested perception an object is autonomous, attended to for its own sake: the disinterested stance is characterized by the absence from the
scene of purposes exterior or ulterior to the object.

The mode of attention which is characterized by this purposelessness may be termed 'intransitive' in that the act terminates upon the object of attention without referring to some entity exterior to that object.6

It is in failing to meet the condition of purposelessness that many general forms of perception differ from aesthetic perception. The observer whose eyes strain to value a jewel or to plot Black's next move is not engaged in aesthetic perception. That mode of perception is attained only by banishing from the attention concern for ends outside the object of contemplation itself. Thus, not all perception of a work of art is aesthetic perception: concern for the patriotic emotion evinced by some of the compositions of Edward Elgar or for the accuracy with which some paintings of Joseph Wright depict eighteenth-century scientific experimentation is strictly outside the bounds of aesthetic perception, though it may of course form an important part of a wider description or judgement of the work.7

Elucidation of the nature of aesthetic disinterestedness and examples of the application of the mode of attention characterized by it may be gleamed from the proponents of this theory in philosophical aesthetics. E. Bullough introduces the notion of aesthetic disinterestedness by considering the perception by an observer of a sea-fog. He notes that for most observers under ordinary conditions, and certainly for mariners, a fog at sea is the cause of a sensation of acute unease or discomfort. The reason for which these sensations may arise in the observers is that sailors or passengers of a ship caught in fog do not generally turn upon their surroundings an attention shorn of utilitarian interest: they are likely to be in fear for the craft's safety or for their lives.8 The mode of attention which they most naturally exercise is thus coloured by consideration of their interests in the matter of the navigation of their vessel.

6 'Intransitivity' is the term introduced by Vivas (1955), pp. 93-9, in defending the tenet that disinterested attention is characteristic of aesthetic perception.
7 Outlined in this paragraph is a thesis about aesthetic perception, which rules that all attention for utilitarian concerns falls outside that category. Related to this thesis is the more extreme aestheticist doctrine about works of art which holds that such objects do not admit of utilitarian or non-aesthetic ends. (The aestheticist doctrine, which has found expression in Cousin's phrase 'art for art's sake', is more extreme than the theory of aesthetic disinterestedness developed in this section in that the latter admits that one may perceive utilitarian ends in a work of art, albeit not through an act of aesthetic perception, whereas the former denies this possibility.) As a thesis about works of art rather than the mode of aesthetic perception, aestheticism falls outside the bounds of the present treatment.
8 The force of this point is intensified by the recollection that the Titanic was lost in fog in the year in which Bullough's article was published.
scene of purposes exterior or ulterior to the object.

The mode of attention which is characterized by this purposelessness may be termed 'intransitive' in that the act terminates upon the object of attention without referring to some entity exterior to that object.\(^6\)

It is in failing to meet the condition of purposelessness that many general forms of perception differ from aesthetic perception. The observer whose eyes strain to value a jewel or to plot Black’s next move is not engaged in aesthetic perception. That mode of perception is attained only by banishing from the attention concern for ends outside the object of contemplation itself. Thus, not all perception of a work of art is aesthetic perception: concern for the patriotic emotion evinced by some of the compositions of Edward Elgar or for the accuracy with which some paintings of Joseph Wright depict eighteenth-century scientific experimentation is strictly outside the bounds of aesthetic perception, though it may of course form an important part of a wider description or judgement of the work.\(^7\)

Elucidation of the nature of aesthetic disinterestedness and examples of the application of the mode of attention characterized by it may be gleamed from the proponents of this theory in philosophical aesthetics. E. Bullough introduces the notion of aesthetic disinterestedness by considering the perception by an observer of a sea-fog. He notes that for most observers under ordinary conditions, and certainly for mariners, a fog at sea is the cause of a sensation of acute unease or discomfort. The reason for which these sensations may arise in the observers is that sailors or passengers of a ship caught in fog do not generally turn upon their surroundings an attention shorn of utilitarian interest: they are likely to be in fear for the craft’s safety or for their lives.\(^8\) The mode of attention which they most naturally exercise is thus coloured by consideration of their interests in the matter of the navigation of their vessel.

---

\(^6\) 'Intransitivity' is the term introduced by Vivas (1955), pp. 93-9, in defending the tenet that disinterested attention is characteristic of aesthetic perception.

\(^7\) Outlined in this paragraph is a thesis about aesthetic perception, which rules that all attention for utilitarian concerns falls outside that category. Related to this thesis is the more extreme aestheticist doctrine about works of art which holds that such objects do not admit of utilitarian or non-aesthetic ends. (The aestheticist doctrine, which has found expression in Cousin's phrase 'art for art's sake', is more extreme than the theory of aesthetic disinterestedness developed in this section in that the latter admits that one may perceive utilitarian ends in a work of art, albeit not through an act of aesthetic perception, whereas the former denies this possibility.) As a thesis about works of art rather than the mode of aesthetic perception, aestheticism falls outside the bounds of the present treatment.

\(^8\) The force of this point is intensified by the recollection that the Titanic was lost in fog in the year in which Bullough's article was published.
`Nevertheless', Bullough writes, `a fog at sea can be a source of intense relish and enjoyment. Abstract from the experience of the sea fog, for the moment [...] direct the attention to the features "objectively" constituting the phenomenon.' Bullough suggests that, if this effort is made, the attention of the observer is absorbed by the fog's milky veil, the carrying power of the air, the creamy smoothness of the water, the feeling of remoteness from the world. When perception is directed to such features for their own sake rather than out of concern for the dangers which the atmospheric conditions pose to navigation, `the experience may acquire [...] a flavour of such concentrated poignancy and delight as to contrast sharply with the blind and distempered anxiety of its other aspects.'

This difference of outlook is due to the observer's dissociating from the utilitarian interests which he or she holds in the situation, and attempting to exercise a mode of disinterested attention. Bullough speaks of this abstraction as the intervention of a measure of `psychical distance' between the observer and the objects of contemplation. `Distance [...] is obtained by separating the object and its appeal from one's own self, by putting it out of gear with practical needs and ends. Thereby the "contemplation" of the object becomes alone possible.'

Bullough characterizes aesthetic perception as that mode of perception which interposes such psychical distance between the observer and his or her objects, which puts the object of perception `out of gear' with the observer's own interests in the object. It is the interposition of distance which constitutes a characteristic feature of aesthetic perception: aesthetic perception is concerned with only the non-utilitarian values in the objects. Thus, a distinction may in Bullough's view be drawn between aesthetic values which are non-utilitarian and several categories of utilitarian values, which include scientific and ethical values.

It is Distance which makes the aesthetic object `an end in itself.' [...]
In particular, it is Distance, which supplies one of the special criteria of aesthetic values as distinct from practical (utilitarian), scientific, or social (ethical) values. All these are concrete values, either directly personal as utilitarian, or indirectly remotely personal, as moral values.

Bullough thus raises the view that there may broadly speaking exist two modes of perception, which differ according to the stance which they maintain in regard

---

9 Bullough (1912), p. 88.
10 Ibid., p. 89.
11 Ibid., p. 91.
12 Ibid., pp. 117-8; emphasis in the original. By `scientific values' Bullough presumably means the empirical power enshrined in scientific theories which permits inter alia technological application.
CHAPTER FOUR: THE AESTHETIC STANCE

to the possible utilitarian values of the objects: one mode is the aesthetic, which abstracts from such values to consider solely the intrinsic merits of the objects (which thus come to be defined as the ‘aesthetic values’ of those objects) while the other mode is the practical, under which the attention is directed to one of the several categories of utilitarian values which an object of perception may possess. One and the same object is clearly susceptible to being perceived under both modes, though presumably not simultaneously; one thus envisages attributing to an object of perception a degree of aesthetic value conferred in the course of disinterested perception, and separately a degree of practical value – in one of the forms that this value can assume – awarded in the act of utilitarian perception.¹³

Bullough’s treatment of the notion of aesthetic perception constitutes an influential contribution to twentieth-century philosophical aesthetics. It has naturally not remained immune from criticism, some of which easily finds justification. One of the chief weaknesses of Bullough’s theory is the presumption that there must exist a special kind of act or state of mind, the specific function of which is to make the aesthetic features of objects accessible to the peripient. Most subsequent theorists of the aesthetic attitude have considered the postulation of these special acts and states of mind as both gratuitous and unpalatable on considerations of psychology; these theorists have attributed the role which Bullough gives to special mental acts instead to ordinary mental attributes functioning in special or perceptually disinterested ways.¹⁴

The conception of the aesthetic attitude as an ordinary action of attending which is performed disinterestedly has more recently been pursued in the work of J. Stolnitz, who has taken a leading role in promulgating this theory. Stolnitz characterizes the ‘aesthetic attitude’ as ‘disinterested and sympathetic attention to and contemplation of any object of awareness whatever, for its own sake alone.’¹⁵ On this characterization, as on Bullough’s, aesthetic perception is performed with ‘no concern for any ulterior purpose’.¹⁶

While Stolnitz’s formulation of the theory of aesthetic disinterestedness is more sophisticated and satisfying than that of Bullough, a well-disposed but moderate reader is likely to conclude that in one respect Stolnitz pursues with

---

¹³ The juxtaposition of two modes of perception, aesthetic and utilitarian, envisaged by Bullough will come to resemble the distinction between two ways of regarding and evaluating scientific theories which will be pursued in later chapters.

¹⁴ For criticism of Bullough’s theory of disinterested attention – including an attack on his notion of special aesthetic states of mind – see Dickie (1974), pp. 91-112.

¹⁵ Stolnitz (1960), pp. 34-5.

¹⁶ Ibid., p. 35.
clearly excessive zeal the thesis that aesthetic perception may pay regard to no end ulterior to the object of perception, and reaches a conclusion which conflicts with most intuitive notions of aesthetic perception. Stolnitz identifies as one of the ends which transcend the object of perception – and therefore an end the pursuit of which is barred to perception which is fully aesthetic – the aim of applying evaluative categories to the object of perception. In his words, if the percipient of a work of art ‘has the purpose of passing judgement upon it, his attitude is not aesthetic.” This statement appears to blur the important distinction between evaluations of an object of perception as a means to the accomplishment of an end ulterior to itself (for instance, the evaluation of a piece of literature as a vehicle for moral reform) and the evaluation of the intrinsic structure or significance of the object with no regard for concerns exterior to it. Evaluations of the former kind are, on the definition of aesthetic perception elaborated by the tradition of Bullough and Stolnitz, foreign to truly aesthetic perception, since such evaluations refer essentially to ends exterior to the object and to the aptitude of the object to attain them. But evaluations of the second kind pay no regard to such exterior ends, and are concerned with what common aesthetic discourse considers to be paradigmatically aesthetic values. It therefore seems contrary to the spirit of Stolnitz’s own approach to rule out the possibility of evaluations of this kind in the exercise of aesthetic perception.18

The present treatment will draw heavily upon Bullough’s and Stolnitz’s conceptions of aesthetic perception as disinterested, but will not pursue the further step envisaged by Stolnitz and deny the possibility of aesthetic evaluation of objects of perception; indeed, much of what follows will constitute a study of the evaluative categories applied by percipients to objects of aesthetic attention.

3. The rise of the discipline of aesthetics

While the theory that disinterested attention is a characterizing feature of aesthetic perception has attracted much comment in the twentieth century, its original

17 Ibid.
18 The view aired by Stolnitz that evaluation of an object of perception cannot be performed as part of an act of aesthetic perception is further criticized in Dickie (1974), pp. 127-8.
CHAPTER FOUR: THE AESTHETIC STANCE

formulation dates from a time 200 years earlier. The present section will briefly retrace the historical origin of this aesthetic theory.

Until the seventeenth century the rules for producing and for evaluating works of art were customarily supported by appeal to the works of Aristotle or the model of classical writers and artists. The late Renaissance disturbed this reliance on authority with two lines of philosophical inquiry which reached their height in the Enlightenment. One current was inspired from Cartesian rationalism and culminated in the mid-eighteenth century in the works of A.G. Baumgarten and G.E. Lessing. This school was animated by the hope that the rules for producing and judging art could be attributed a solid, a priori foundation by deduction from basic, self-evident axioms, such as the principle that art is the imitation of nature.

The other line was pursued principally by English and Scottish empiricist writers. Rather than in the rigorous foundation of aesthetic canons of appraisal, these were interested in the psychology of art and especially in the creative process and the effect of art on the beholder. Emphasis in the theory of art shifted in this tradition from object to subject, from the work of art to its perceiver and critic, shadowing the wider transfer of emphasis by the English and Scottish Enlightenment (apparent in, for instance, Humean epistemology) from the rational to the sensate.

It is in the work of this school that many elements of the modern discipline of aesthetics were forged. The first task of aesthetics was to constitute its own discipline: it had to be demonstrated that there existed a body of experience not accounted for by other philosophical disciplines and which could therefore constitute the subject-matter of a distinctive discipline of aesthetics. The identification of a peculiar aesthetic approach or attitude to philosophical objects was required.

This search was concluded by the insight that while particular philosophical disciplines studied the act of perception of objects in the pursuit of given external values or concerns, as yet none of these disciplines studied pure perception or the act of perception detached from ulterior concerns. It is in this way that the notion that disinterestedness is constitutive of aesthetic perception

19 On the history of aesthetics from the Renaissance to the Enlightenment see e.g. Kristeller (1952); of that essay pp. 17-24, 32-43 deal with the Cartesian tradition, pp. 24-31 with the empiricist.

20 The claim that the British empiricist tradition made in this way an important contribution to the foundation of the discipline of aesthetics is supported in e.g. Stolnitz (1961b), pp. 131-2. Cassirer believes that Shaftesbury founded 'the first really comprehensive and independent philosophy of the beautiful' (1932, p. 312).
was coagulated. The use of the term 'aesthetics' to designate a field of study was introduced by Baumgarten; but it was in many ways the empiricist school that most fully attributed to aesthetics its etymological sense of the pure study of perception or contemplation. It is easy to see how such a stance is apt to promote the autonomy not only of philosophical aesthetics but also of art criticism: a work of art must now be evaluated on its perceptual merits rather than as moral or cognitive vehicle.\(^{21}\)

More specifically, the notion of disinterestedness came into prominence largely in opposition to two pre-existing philosophical doctrines: the rhetorical tradition's view of art-works as instruments to 'instruct by pleasing', and the 'intelligent egoism' of T. Hobbes, who had argued that the precepts of morality and religion can be retraced to enlightened self-interest. The first philosopher to call attention to the notion of disinterested perception was the third Earl of Shaftesbury: his writings on this topic mirror the progressive demarcation of aesthetic from other philosophical speculation, as the notion of disinterestedness originates in his ethical doctrines but soon acquires a purely aesthetic import.\(^{22}\)

Shaftesbury aimed to demonstrate against Hobbes that it is natural to man to be unselfish. While Hobbes reconstructed the dictates of morality as the codification of a consistent pursuit of self-interest, Shaftesbury insists that 'there can be nothing more fatal to virtue than the weak and uncertain belief of a future reward and punishment.'\(^{23}\) The only legitimate motivation for a man's undertaking what is morally righteous is his apprehension of its righteousness:

\[
\text{Though he may intend to be virtuous, he is not become so for having only intended or aimed at it through love of the reward. But as soon as he is come to have any affection towards what is morally good, and can like or affect such good for its own sake, as good and amiable in itself, then he is in some degree good and virtuous, and not till then.}^{24}\]

In cautioning against self-interest as an apt motivation for the pursuit of virtue Shaftesbury thus urges in its place not altruism, the common good or some other

\(^{21}\) Stolnitz (1961a, p. 100) argues that the notion of aesthetic disinterestedness is not to be found in classical or medieval aesthetic writing and is thus genuinely new to the eighteenth century. The thesis that disinterestedness was a determining principle in the rise of aesthetics in the Enlightenment is forcefully argued by Stolnitz (1961b), on which the present treatment draws.

\(^{22}\) Further on Shaftesbury's aesthetics in general see Brett (1951), pp. 123-44; on his notion of disinterestedness and its ethical forebears see Stolnitz (1961a), pp. 105-13. On the relation between ethical and aesthetic values in Shaftesbury see J.A. Bernstein (1980), pp. 21-60.

\(^{23}\) Shaftesbury (1711), vol. I, p. 275.

\(^{24}\) Ibid., pp. 273-4.
desirable consequence, but disinterestedness towards all ends.

The theory of disinterestedness thus originates in the practical domain, for it refers to (or, more determinately, rejects as unsuitable to the attainment of virtue) actions directed towards certain goals; but the notion acquires a perceptual import when Shaftesbury turns to explain how the moral agent can recognize and attain rectitude. Virtue, states Shaftesbury, exhibits a harmony or propriety which can be apprehended by the agent as perceiver. Shaftesbury explicitly likens this perceptual operation to the appreciation of beauty in material objects:

The case is the same in the mental or moral subjects as in the ordinary bodies or common subjects of sense. The shapes, motions, colours, and proportions of these latter being presented to our eye, there necessarily results a beauty or deformity, according to the different measure, arrangement, and disposition of their several parts. So in behaviour and actions [...]. The mind [...] finds a foul and fair, a harmonious and a dissonant, as really and truly here as in any musical numbers or in the outward forms or representations of sensible things.25

In Shaftesbury the virtuous agent is conceived as a perceptual subject devoted to 'the very survey and contemplation' of beauty in morals. While 'disinterestedness' thus initially refers to an attitude in moral conduct, elaboration of the notion elucidates its perceptual import. Moral rectitude is attained not through the cultivation of interests but by dissociating one's conduct from all interests and searching for perceptual propriety in moral acts: the attitude of a 'disregard for interests' thus comes to characterize the mode of attention exercised in aesthetic perception.

Shaftesbury's most extended discussion of aesthetic disinterestedness occurs in one of his later works in which the perceptual commands its own attention divorced from moral considerations. Shaftesbury characterizes the attitude of aesthetic disinterestedness in contradistinction to a perceptual stance animated by a desire to command or use an object of perception to some purpose. Aesthetic contemplation, or the apprehension of beauty, attributes to the object of perception no features not intrinsic to the very act of perception, and certainly no utilitarian dimension. Shaftesbury suggests that such non-utilitarian perception is

---

25 Ibid., p. 251. Hutcheson, Shaftesbury's follower whose views on aesthetic disinterestedness will be discussed below, pursues a similar analogy between the apprehension of moral propriety and the perception of natural beauty: 'The Author of nature has much better furnished us for a virtuous conduct than some moralists seem to imagine, by almost as quick and powerful instructions as we have for the preservation of our bodies. He has given us strong affections to be the springs of each virtuous action, and made virtue a lovely form, that we might easily distinguish it from its contrary, and be made happy by the pursuit of it' (1725, p. 25).

55
CHAPTER FOUR: THE AESTHETIC STANCE

commonly directed upon elements of the natural world which evade human manipulation:

Imagine [...] if being taken with the beauty of the ocean, which you see yonder at a distance, it should come into your head to seek how to command it, and, like some mighty admiral, ride master of the sea, would not the fancy be a little absurd? 26

The fulfilment which would derive from ‘possessing’ the ocean is ‘very different from that which should naturally follow from the contemplation of the ocean’s beauty.’ 27

It was remarked in the previous chapter that the characterization of aesthetics by reference to a certain mode of perception rather than to a stipulated class of ‘aesthetic objects’ facilitated the extension of aesthetic categories to intellectual constructs, and thereby fostered the study of the aesthetic aspects of scientific theories. Shaftesbury illustrates the facility with which theories of aesthetics of this form can approach the notion of the beauty of intellectual objects when he turns from the perception of beauty in nature to the aesthetic delight afforded by contemplation of theorems in mathematics:

There is no one who, by the least progress in science or learning, has come to know barely the principles of mathematics, but has found, that in the exercise of his mind on the discoveries he there makes, though merely of speculative truths, he receives a pleasure and delight superior to that of sense. When we have thoroughly searched into the nature of this contemplative delight, we shall find it of a kind which relates not in the least to any private interest of the creature, nor has for its object any self-good or advantage of the private system. 28

Shaftesbury’s remarks on mathematics appear to foreshadow those of Bullough on ‘scientific values’ discussed in the previous section: the formulations of both these theorists of aesthetic disinterestedness raise the prospect of the existence of two modes of attention which may be directed upon the products of scientific

26 Shaftesbury (1711), vol. II, p. 126. Shaftesbury also rebuts the objections of those who argue that one experiences art for the sake of the pleasure which will derive, by means of a distinction between the aesthetic appreciation of beauty and the second-order (non-aesthetic) awareness of the gratification thereby produced: ‘Though the reflected joy or pleasure which arises from the notice of this pleasure once perceived, may be interpreted a self-passion or interested regard, yet the original satisfaction can be no other than what results from the love of truth, proportion, order and symmetry in the things without’ (ibid., vol. I, p. 296). Hutcheson draws a similar distinction: ‘It plainly appears that some objects are immediately the occasions of this pleasure of beauty, and that we have senses fitted for perceiving it, and that it is distinct from that joy which arises upon prospect of advantage’ (1725, p. 37).


CHAPTER FOUR: THE AESTHETIC STANCE

research. One, implicitly allowed in the passage of Shaftesbury reproduced above, relates to the interests which the percipient may hold in the use of a theorem or theory; the other, which Shaftesbury is concerned to distinguish from the former, yields a ‘contemplative delight’ which is wholly disjoint from prospects of utilitarian advantage. Again, the ideas underlying this partition of concerns between modes of attention will be exploited later in the present work to suggest that scientific practice commonly evaluates theories separately on a logico-empirical (and thus utilitarian) canon on the one hand and on an aesthetic (or disinterested) canon on the other.

The aesthetic works of Hutcheson were examined in the previous chapter for their references to the application of aesthetic categories to constructs of the intellect. It was suggested there that his readiness to extend aesthetic discourse to intellectual objects was prompted by his general aesthetic theory, but illustration of this connection has thus far been postponed. Hutcheson’s views on aesthetics in general follow closely those of Shaftesbury, and indeed the treatise containing almost all his writings on aesthetics was conceived as a defence of the theories of his predecessor. Hutcheson too suggests that the verdicts delivered by aesthetic judgements are independent of any thought for the consequences of the act of judgement:

Neither can any resolution of our own, nor any prospect of advantage or disadvantage, vary the beauty or deformity of an object. [...] So propose the whole world as a reward, or threaten the greatest evil, to make us approve a deformed object, or disapprove a beautiful one: dissimulation may be procured by rewards or threatenings, or we may in external conduct abstain from any pursuit of the beautiful, and pursue the deformed, but our sentiments of the forms, and our perceptions, would continue invariably the same. [...] Nay, do not we often see convenience and use neglected to obtain beauty, without any other prospect of advantage in the beautiful form than the suggesting the pleasant ideas of beauty? 29

On this basis Hutcheson excludes from the aesthetic stance any concern for utilitarian benefit. Such profit may be welcomed and may arouse pleasure; but this pleasure is wholly different from the appreciation of beauty. It follows that the aesthetic interest is indifferent to the causal and other relationships which the object has to things beyond itself.

Under this conception, nothing is a priori barred from the scope of aesthetic

---

29 Hutcheson (1725), pp. 36-7. Hutcheson’s question, whether we do not often ‘see convenience and use neglected to obtain beauty’, with ‘no prospect of advantage’ save ‘the pleasant’, will assume relevance to theory-choice in science in a later chapter where it will be suggested that adherents of degenerating paradigms in scientific revolution pursue beauty in theories to the neglect of empirical virtues.
criticism: it becomes an empirical question whether aesthetic interest is aroused by a given class of entities. Like Shaftesbury before him, Hutcheson thus considers the feelings aroused by the contemplation of a scientific theory and concludes that such objects too ought to fall within the domain of aesthetic discussion.

In particular he distinguishes the aesthetic contemplation of a theory from an awareness of the narrowly scientific or empirical utility which its application affords:

It is easy to see how men are charmed with the beauty of such knowledge, besides its usefulness [...]. And this pleasure we enjoy even when we have no prospect of obtaining any other advantage from such manner of deduction than the immediate pleasure of contemplating the beauty.\(^\text{30}\)

It is hoped that the constant recurrence of the distinction between the aesthetic or disinterested and the empirical or utilitarian modes of perception applicable to intellectual constructs which characterizes the history of this aesthetic theory will prepare the ground for the claim in the next chapter that scientific practice similarly distinguishes between empirical and aesthetic evaluations of scientific theories.

Aided by the great influence enjoyed by the works of Shaftesbury and Hutcheson, the notion of aesthetic disinterestedness dominated much of the remainder of in eighteenth-century British philosophy, finding for instance in A. Gerard one of its proponents in the latter half of the century.\(^\text{31}\) But it was I. Kant who reformulated and lent rigour to the ideas less systematically discussed by the eighteenth-century English aestheticians.\(^\text{32}\)

Kant brought the theory of aesthetic judgement within the ambit of a general theory of ‘teleological judgement’ by which he believed that he had bridged the gulf excavated in his previous works between the concept of Nature, which was the realm of law and science, and the concept of Freedom, which was the realm of self-imposed rational principles or ends. In the teleological judgement, under which he subsumes the aesthetic judgement, he believed that he had found the link between the spheres of natural science and morality.

In the opening sections of the ‘Critique of Aesthetic Judgement’ Kant casts

\(^{30}\) Ibid., p. 51.


\(^{32}\) On disinterestedness and attitude in Kant’s aesthetics see Osborne (1968), pp. 113-30, and McCloskey (1987), pp. 29-49, on which I have drawn.
into a logical mould Shaftesbury’s discrimination of a specifically ‘aesthetic’ attitude to the world. Where the English aestheticians had been interested in the psychological description of the aesthetic attitude and the psychological differences between it and practical attitudes, Kant concerned himself with the grounds of judgement, differentiating the logical basis of aesthetic judgements from that of the judgements we make about the other kinds of pleasure which things give us, and from that of judgements about utility and goodness.

To differentiate aesthetic judgements from moral judgements, judgements of utility or judgements of sensual pleasure, Kant turns to Shaftesbury’s notion of disinterested attention. Kant believed that aesthetic judgements must be entirely free of interest. ‘Every one must allow that a judgement on the beautiful which is tinged with the slightest interest, is very partial and not a pure judgement of taste.’ Coherently with this insistence, the First Moment of the Analytic of the Beautiful defines the the faculty of aesthetic perception and the quality of the object of perception capable of yielding aesthetic pleasure as follows: ‘Taste is the faculty of estimating an object or a mode of representation by means of a delight or aversion apart from any interest. The object of such a delight is called beautiful.’

With the influence of Kant the notion of aesthetic disinterestedness attained a central position in aesthetic speculation. In particular, the idea that disinterested pleasure was distinctive of aesthetic experience remained central to Schiller and the German Idealists.

33 Kant (1790), p. 43.
34 Ibid., p. 50; emphasis in the original. The English aestheticians had contrasted the disinterested attitude with ‘interest’ in the sense of a desire to possess or manipulate a thing. Kant excluded from the aesthetic attitude not merely considerations of advantage but any concern for the very existence of a thing. To reach a pure aesthetic experience, he writes, ‘one must not be in the least prepossessed in favour of the real existence of the thing, but must preserve complete indifference in this respect, in order to play the part of judge in matters of taste’ (ibid., p. 43).
Chapter Five
THEORY-EVALUATION ON AESTHETIC GROUNDS

One can always make a theory, many theories, to account for known facts, occasionally even to predict new ones. The test is aesthetic.¹

Sir George Thomson

1. A conceptual vacancy is announced and filled

Here the concerns and approaches of the foregoing chapters will be conjoined to yield the conclusion which will engage the central section of the present treatment. Chapter 2 surmised that reference to logico-empirical evaluative criteria was alone insufficient to reconstruct salient instances of theory-evaluation, and that recourse to extra-empirical criteria was required. In an apparent break with the concerns of the philosophy of science, chapter 3 suggested the existence in scientific and wider cultural circles of a practice of discerning aesthetic value in intellectual constructs, and retraced some of the philosophical ancestry of the application of aesthetic categories to such objects. Chapter 4 explored the theory central to Western philosophy and art criticism that aesthetic perception is characterized by disinterested attention; this strand of aesthetic theory views perception as an act detached from considerations of the practical utility or advantage afforded by the object of contemplation.

The several investigations in the historiography and methodology of science and in the theory of aesthetic perception here briefly summarized bring to mind an intuitively attractive synthesis. Logico-empirical evaluative criteria prize those theoretical qualities which are held in methodology to be constitutive of observational success, the posited goal of science. The historian's intuition that there must sit over theories a jury alongside that composed of logico-empirical criteria opens a conceptual vacancy: since its announcement in chapter 2 above, the search has been joined for a class of non-empirical evaluative criteria for use in the reconstruction of the history of science and in characterizations of scientific

rationality.

The terms of the search stipulate some requirements to be satisfied by any class of criteria which aspires to fill the vacancy: one wishes for a class of criteria of theory-evaluation which are disjoint from logico-empirical evaluative criteria in being unconcerned for the empirical success of the theories to which they are applied in judgement. One should in addition like prima facie evidence to show that such criteria have been widely applied in significant instances of theory-evaluation and choice in history; one may lastly hope that the class of criteria which meets these requirements be philosophically interesting, in that it is itself susceptible to substantial philosophical analysis or that it contributes to elucidating the nature of scientific rationality.

While chapter 2 announced the vacancy and urged a search for a class of non-empirical evaluative criteria by which it might be filled, the two subsequent chapters laid the grounds upon which such a class may be constructed. The present treatment will propose that the vacancy be filled by a canon of aesthetic evaluative criteria which embody the disinterested perceptual attitude specified in chapter 4. Since this attitude is characterized by unconcern for forms of utility or advantage, the canon of evaluative criteria which embodies it will fulfill to the highest standard the requirement of extra-empiricality: this canon will remain aloof from calculations of the technological applicability demonstrated, the predictive power exhibited or even – the least nakedly utilitarian of the logico-empirical concerns of scientists – the observational success attained by the scientific theories which constitute its objects of perception. This second canon will evaluate theories by reference to qualities quite unlike those – which are constitutive of observational success – of predictive accuracy or the generation of novel predictions: it will instead value non-utilitarian, aesthetic qualities of theories.

The suggestion advanced here is thus that in scientific practice there is a second jury sitting over theories beside that of the logico-empirical criteria, evaluating theories by reference to aesthetic criteria. A scientific community’s overall canon of theory-assessment – the effects of the application of which to instances of theory-choice transpire in the history of science – is composed of the two classes of logico-empirical and aesthetic evaluative criteria. Although, as will be argued in later chapters, the two classes of evaluative criteria differ in many respects including the manner of their historical origin and the degree of their temporal variability, the present treatment intends to attribute to them equal prima facie methodological dignity, in an even-handedness which demands that
serious philosophical attention be directed at the aesthetic evaluative canon no less than to the more familiar logical-empirical canon.

In the intention of the present treatment, this suggestion will itself be judged on the two grounds outlined in the methodological introduction supplied in chapter 2: the reader will be invited to test the facility with which a philosophical view of science incorporating this suggestion can provide a coherent and convincing reconstruction of past science, and separately to gauge the degree of the suggestion's philosophical fertility. These two tests will be applied to the present suggestion in chapters to come.

In the matter of philosophical fertility, the suggestion that the historian's intuition be married in this way to the aesthetic theorist's insight holds much that is new. For instance, chapter 7 will use the nature of aesthetic canons of theory-evaluation to suggest a mechanism by which elements of scientific methodology are constructed by scientific communities from historiographical sources, and chapter 8 will present an interpretation of the notion of scientific revolution as aesthetic rupture. It will however prove necessary to establish much of this view in the teeth of entrenched opposition from conventional schools of philosophy of science. While perhaps few writers deny to aesthetic considerations any occurrence whatever in science, many attribute to them a severely circumscribed role in the expectation that their subsequent neglect will prove justified. It will be necessary to demonstrate them wrong.

In the matter of the power of this suggestion to help account for the history of science, detailed evidence will be marshalled below. It suffices here to advance some prima facie evidence that there exists a practice of the aesthetic evaluation of theories by scientists. While the pronouncements which follow cannot be taken as authoritative descriptions of the practice of scientific communities at large, they convey a measure of the importance which is apparently attributed to aesthetic theory-evaluation by notable members of such communities. G.H. Hardy has for instance written:

>The mathematician's patterns, like the painter's or the poet's, must be beautiful; the ideas, like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics.\(^2\)

P.A.M. Dirac stipulates for theories a similar aesthetic requirement:

---

\(^2\) Hardy (1940), p. 85; emphasis in the original. It would not be pertinent to object to the relevance of this quotation on the grounds that mathematics cannot be counted amongst the sciences: the mathematical structure of modern physical theory tends to conflate the notions of mathematical and scientific aesthetics.
CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

It is more important to have beauty in one's equations than to have them fit experiment. [...] It seems that if one is working from the point of view of getting beauty in one's equations, and if one has really a sound insight, one is on a sure line of progress.³

Lastly A. Einstein was thus depicted by his son Hans Albert, himself a physicist:

He had a character more like that of an artist than of a scientist as we usually think of them. For instance, the highest praise for a good theory or a good piece of work was not that it was correct nor that it was exact but that it was beautiful.⁴

It will be noted that the subjects of these passages clearly elevate the requirement that theories should possess certain (as yet unspecified) aesthetic features to the status of criteria by which they explicitly propose to evaluate the theories which come before them. The above testimony also distinguishes unequivocally between on the one hand the cognitive or empirical qualities exhibited by theories – qualities summarized by Dirac as 'experimental fit' and by Einstein as 'correctness and exactitude' – and on the other hand the theories' aesthetic features. Thus Dirac distinguishes between 'having beauty in one's equations' and 'having them fit experiment'; Einstein between 'a correct theory' and 'a beautiful one'. This distinction is here taken as preliminary confirmation that the scientific practice which makes use of aesthetic criteria in theory-evaluation sees them as quite distinct from the logico-empirical evaluative criteria which may simultaneously be applied to the same theories, and in particular that the possession of aesthetic features by scientific theories is not held to be constitutive of their empirical success, as logico-empirical features on the contrary are.

The above statements contain clear echoes of the distinction promulgated by the aesthetic theory of disinterested attention between the utilitarian and the aesthetic modes of attention. These echoes are gladly welcomed by the present

³ Dirac (1963), p. 47. Throughout his career Dirac attributed an important role to aesthetics in scientific research; reminiscences of and comments on his opinions on this matter appear in Kursunoglu and Wigner eds (1987) and J.G. Taylor ed. (1987), passim. Dalitz (1987, p. 20) for instance recounts an episode during Dirac's tenure of a visiting professorship at the University of Moscow in 1955: 'When asked to write briefly his philosophy of physics, he wrote on the blackboard “PHYSICAL LAWS SHOULD HAVE MATHEMATICAL BEAUTY” and this has been preserved there to this day.' Dirac's most notable application of aesthetic criteria, in the evaluation of the theory of quantum electrodynamics, is referred to in section 4 of the present chapter.

⁴ Whitrow ed. (1967), p. 19. This character-sketch is corroborated by Bondi ibid., p. 82: 'As soon as an equation seemed to him to be ugly, he really rather lost interest in it [...]'. He was quite convinced that beauty was a guiding principle in the search for important results in theoretical physics.' Dirac appears more insistent: 'Einstein seemed to feel that beauty in the mathematical foundation was more important, in a very fundamental way, than getting agreement with observation' (1982, p. 83).
treatment and are here taken as initial vindication of the suggestion that the vacancy left open by scientific methodology be filled by aesthetic theory.\(^5\)

The remainder of the present treatment will assume that there exist and are widely applied in scientific practice such entities as aesthetic canons of theory-appraisal, the nature, origin and behaviour of which will be elucidated progressively. An aesthetic canon will be assumed to be composed of a number of 'aesthetic criteria'; each criterion consists of a statement attaching a certain preferential weighting to an aesthetic feature or quality apt to be exhibited by scientific theories. To apply the canon in an instance of theory-appraisal one ascertains the aesthetic features of the theory in question and determines what preferential weighting is attached to them by the evaluative canon. The weightings which the canon attaches to the aesthetic features of a given theory constitute a measure of the degree of preference extended by the canon to that theory; in a case of the comparative evaluation of several theories, the theory which obtained the greatest preference would be deemed to have been – so to speak – nominated by the aesthetic canon for endorsement by the scientist or community performing the evaluation. (Whether that theory will at the end of the day be chosen for endorsement by the agent depends in part also of course on the verdicts independently delivered on the competing theories by the logico-empirical evaluative canon.)

2. Positivist denials of aesthetic theory-evaluation

That the above remarks would raise the hostility of much traditional philosophy of science is clear. This section and the next present and rebut some influential initial criticism which would be moved against the enterprise sketched out above

\(^5\) The above couple of paragraphs have briefly mustered evidence for the proposition that scientific practice makes use of aesthetic criteria of theory-evaluation, and furthermore that such criteria are disjoint from the class of logico-empirical criteria. The evidence adduced in the main text was culled from the words of practising scientists themselves, but clearly some such evidence is transmitted by the mediation of works of historiography. Thus Neyman (1974, p. 9) writes that 'Copernicus introduced a completely novel yardstick for appraising a new theory: conformity with observations and intellectual elegance.' Neyman's remark appears to attribute to Copernicus's work the same distinction between logico-empirical and aesthetic evaluative canons enshrined in the statements of twentieth-century scientists cited above. The incidence of aesthetic features and considerations in the work of Copernicus will further be studied in chapter 10.
by positivist philosophy. While the present section outlines the thrust of this criticism, the next attempts to respond to it; this manoeuvre will enable some further evidence to be presented in favour of the claims made in section 1 above.

Positivism established a demarcation between two stages of scientific procedure which boasted prima facie plausibility for several decades. It distinguished between an arational context of discovery in which theory-generation occurred by enigmatic intuitions or conjectures and a context of justification where inference from empirical data tested the products of the former stage and assured the rationality of science. The distinction was voiced by e.g. R.S. Rudner:

In general, the context of validation is the context of our concern when, regardless of how we have come to discover or entertain a scientific hypothesis or theory, we raise questions about accepting or rejecting it. To the context of discovery, on the other hand, belong such questions as how, in fact, one comes to latch on to good hypotheses, or what social, psychological, political, or economic conditions will conduce to thinking up fruitful hypotheses.

A bare statement of the distinction between discovery and justification might lead a student unappraised of the literature to expect its aim to have been to identify the process of discovery as singularly interesting: after all, on positivist grounds that process appears to await intricate explanation referring to the circumstances and experiences of individual scientists. No explanations are by contrast seemingly required by the positivist in the context of justification: a mere exposition of the empirical facts would seem sufficient to entail (thus satisfying the strictest positivist standards of explanation) the acceptance or rejection of theories. As the most cursory reading of positivist literature will persuade, the intention of the discovery/justification distinction was in reality the diametric opposite: its result was to evacuate the former term of philosophical interest and prescribe analysis and explanation solely of the latter. K.R. Popper is an advocate of both this partition and the attendant apportionment of weight, as demonstrated by a passage from a paragraph characteristically titled Elimination of Psychologism:

The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it.

---

6 This distinction originated with Reichenbach (1938, pp. 6-7), though it was foreshadowed in certain aspects by Herschel (1830, p. 164). The history of the distinction and of criticism of it is retraced by Hoyningen-Huene (1987).

7 Rudner (1966), p. 6. The author uses the term ‘validation’ in place of ‘justification’.

8 The differential allocation of interest was first endorsed by the originator of the modern version of the discovery/justification distinction, Reichenbach (1949, p. 289): ‘The philosopher of science is not much interested in the thought processes which lead to discovery.’

65
There is no such thing as a logical method of having new ideas, or a logical reconstruction of this process. My view may be expressed by saying that every discovery contains ‘an irrational element’, or ‘a creative intuition’, in Bergson’s sense.9

It is by statements of preference of this form that positivism dismisses from rational attention the province of scientific discovery.

Entertaining for one moment the positivist terms of the distinction between discovery and justification, to which camp should aesthetic criteria be consigned? Aesthetic criteria are intuitively – and with some justice – perceived as non-quantitative, non-empirical, imprecise and arational, of the same form as the ‘psychological conditions’ of scientific thought of which Rudner speaks or the ‘creative intuitions’ of Popper. If the positivist demarcation between discovery and justification is embraced, one readily conceives the belief that aesthetic criteria should operate in the context of discovery but not in that of justification. Under this reading of positivist concerns, if Popper forbears from mentioning aesthetic criteria in his treatment of scientific method, that silence may be due to his implicitly confining their influence to theory-generation in which he believes philosophers of science may profess no rational interest.10

The onlooker’s expectation that positivist philosophy of science would first consign references to aesthetic considerations in scientific practice to the category of phenomena of the context of discovery and secondly presume that in such capacity they are not susceptible to rational analysis is confirmed by a study of positivist treatments of the subject. One of the most comprehensive and explicit illustrations of the positivist attitude towards aesthetic considerations is offered by H. Feigl, who endorses the three stages of this progressive denial of their consequence. He first entrenches the conventional partition of his subject-matter:

The distinction [...] between the *historicosociological* narratives, analyses, causal accounts of the origins, developments, conflicts, and *Zeitgeists* of scientific ideas on the one hand, and the *logicomethodological* reconstructions of scientific knowledge claims, on the other hand [...], remains (and, as I see it, *should* remain) at least a most important first approximation if we are to retain even a

---

9 Popper (1959), pp. 31-2.
10 In fact Popper devotes a paragraph to the aesthetic concept of simplicity, but only – as one might expect – to dismiss it. He states that a consideration of elegance has little interest from the point of view of the theory of knowledge: it does not fall within the province of logic, but merely indicates a preference of an *aesthetic or pragmatic* character (ibid., p. 137; emphasis in the original).
CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

minimum of clear thinking in these badly confused matters.\textsuperscript{11}

He initially betrays no predilection for either of the modes of inquiry which this passage delineates, but soon there remains no doubt where he believes the choice philosophical meat to lie:

Consider, just as prime examples, the writings of Galileo and Newton. Their formulations usually came 'after the fact' of their discoveries - experimental or theoretical. This is prototypical of all good philosophy of science. In Newton's formulation of his assumptions (postulates), definitions (explications), and \textit{regulae philosophandi} (precepts), we find one of the truly great masters giving us a rational reconstruction of his theoretical achievements.\textsuperscript{12}

Having characterized what he understands as 'good philosophy of science' Feigl takes pains to bar aesthetic considerations from its domain, confining them to the context of discovery where philosophy has no jurisdiction:

A few words on some misinterpretations stemming from predominant concern with the history and especially the \textit{psychology} of scientific knowledge. In the commendable (but possibly utopian) endeavor to bring the 'two cultures' closer together (or to bridge the 'cleavage in our culture') the more tender-minded thinkers have stressed how much the sciences and the arts have in common. The 'bridges' [...] are passable only in regard to the \textit{psychological} aspects of scientific [...] creation [...]. Certainly, there are esthetic aspects of science [...]. But [...] what is primary in the appraisal of scientific knowledge claims is (at best) secondary in the evaluation of works of art - and vice versa.\textsuperscript{13}

Thus are aesthetic criteria disenfranchised in the positivist view of science. The sole merit one may according to the positivist ascribe to aesthetic intuitions is that of having a valuable heuristic role in the arational process of theory-generation, on a par with dreams and hallucinations.\textsuperscript{14}

\textsuperscript{11} Feigl (1970), p. 4; emphasis in the original. After a so emphatic endorsement of one of the main tenets of logical positivism it is surprising a few pages on to read an abjuration: 'I personally abandoned, long ago, whatever adherence I had to positivistic [...] philosophies of science' (\textit{ibid.}, p. 7).

\textsuperscript{12} \textit{Ibid.}, p. 6; emphasis in the original.

\textsuperscript{13} \textit{Ibid.}, pp. 9-10; emphasis in the original. This article of Feigl appears in the same collection which contains Buchdahl (1970), cited in chapter 2. It is perplexing to find Feigl's conservative, positivist treatment of evaluative criteria and dismissal of aesthetic evaluations alongside Buchdahl's considerably more innovative perspective.

\textsuperscript{14} The heuristic role (as opposed to the evaluational role, which is the main concern of this treatment) of aesthetic criteria in science is further studied in Mamchur (1987). I am grateful to Dr E. Mamchur of the Institute for Philosophy of Science, USSR Academy of Science, for illuminating discussions on this subject.
CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

3. Transcending the positivist demarcation

Fiegl is correct to the extent to which he discerns the operation of aesthetic criteria in the context of discovery. As the following chapters will illustrate, the history of science indeed teems with instances in which aesthetic criteria appear to have directed theory-formulation, at times prevailing in this role over empirical criteria.

One such episode was witnessed in the development of quantum mechanics from 1925, independently achieved by W. Heisenberg first and E. Schrödinger shortly afterwards. The story of this episode told in 1963 by their junior colleague Dirac has become a locus classicus of discussions of beauty in science:

Heisenberg worked keeping close to the experimental evidence about spectra [...]. Schrödinger worked from a more mathematical point of view, trying to find a beautiful theory for describing atomic events [...]. He was able to extend De Broglie's ideas and to get a very beautiful equation, known as Schrödinger's wave equation, for describing atomic processes. Schrödinger got this equation by pure thought, looking for some beautiful generalization of De Broglie's ideas, and not by keeping close to the experimental development of the subject in the way Heisenberg did.15

Popper and Feigl would experience no qualms in accommodating this report in their conception of scientific method as an instance of arational theory-generation, on the understanding that such a categorization divests it of philosophical interest. Perhaps because Dirac's account of the development of quantum mechanics is so easily reducible to the received or positivistic image of science, it is one of the fragments of evidence from the history of scientific practice most promptly quoted in brief or oblique treatments of the role of aesthetics in science.16

Aesthetic criteria are however for some scientists important not only as illustrated above in theory-generation – an activity which in the view of proponents of the discovery/justification distinction is arational – but also in theory-evaluation and choice, which in this view aspire to full rationality. Dirac, who in the passage reproduced above stressed the weight of aesthetic criteria in

15 Dirac (1963), pp. 46-7. Krisch (1987, p. 51) testifies that Dirac believed quite generally in the heuristic role of aesthetics in theory-formulation: 'Dirac stated that, "...the elegance of the formulation was very important in choosing the direction for one's research."'

CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

directing research, showed no hesitation in extending their role equally to the
evaluation of the products of that research, unfavourably rating the non-linear
spinor theory of Heisenberg in 1967 on a blatantly aesthetic canon: 'My main
objection to your work is that I do not think your basic (non-linear field) equation
has sufficient mathematical beauty to be a fundamental equation of physics.' Positivism could judge this comment by an experienced and eminent physicist in
no way other than deranged, so alien is it from positivist precepts of theory-
evaluation. But positivism is thus compelled to some historiographic evasiveness,
dismissing Dirac's pronouncement of 1967 as an aberration while acknowledging
his report of 1963 as typical of the arational process of discovery. The queries
arise whether there is not more unity in scientific reason than positivism will
allow, whether aesthetic criteria have not an integral role in both context of
discovery and of justification, and lastly whether the positivist demarcation
between the two contexts should not be relaxed.

Appeals to aesthetic criteria of theory-evaluation come to light almost
whenever the reasons for the espousal of a consequential theory are ascertained
by historical inquiry. For instance, the historian of science P.G. Bergmann
attributes the diffusion of Einstein's general theory of relativity jointly to
empirical and aesthetic factors:

From a logical point of view, the progress toward general relativity
depended on a number of choices to be made; its eventual adoption,
first by Einstein himself and later by the community of physicists,
depended on the esthetic appeal of the finished theory and on its
confirmation by experiment and observation.

Dirac's perception of the reasons for which to embrace that same theory attributes
equal prominence to its aesthetic virtues:

The foundations of the theory are, I believe, stronger than what one
could get simply from the support of experimental evidence. The
real foundations come from the great beauty of the theory. [...] It is
the essential beauty of the theory which I feel is the real reason for
believing in it.

The balance of the evidence surely makes it invidious to confine the impact of

of Dirac of March 1967.
18 A scientist's general discussion of the role of aesthetic considerations in theory-
evaluation is offered in Chandrasekhar (1987), pp. 59-73.
20 Dirac (1980), p. 10. A further discussion of the aesthetic properties of the theory of
aesthetic criteria to the stage of theory-generation, particularly if the purpose of this move is to neuter their power by appeal to the supposedly arational character of discovery.

Historically-adequate understanding of the evaluational as well as the heuristic roles played by aesthetic concerns points to a limitation of the notion of 'themata' introduced by G. Holton.\textsuperscript{21} Themata are in Holton's account the fundamental principles that motivate a scientist or a scientific community, appearing and re-appearing throughout history. In Holton's view themata can include specific commitments, such as the commitment to the principles of conservation of mass and energy, as well as general epistemological or methodological principles, such as those which value the attainment of forms of beauty or elegance in theories. Holton's inclusion of the latter possibility raises the hope that his treatment may attribute an important role in science to aesthetic criteria. Unfortunately Holton's methodology appears still to be dominated by the logical-empiricist model of science, and particularly by the distinction between the context of discovery and that of justification. Holton envisages themata as operating almost exclusively at the stage of 'science in the making', when the scientist is engaged in formulating his or her ideas. Once the theories are in the public domain and come to be evaluated, Holton's repertory of themata have – or claim – less relevance: they do not appear to participate in theory-appraisal. By limiting the operation of themata to the phase of discovery, Holton gratuitously barred his own way to a potentially incisive and novel analysis of theory-evaluation.

The limitation imposed by Holton on the domain of operation of themata appears particularly regrettable in those themata which are constituted by aesthetic concerns, for – as the present treatment hopes to establish – the consideration of aesthetic evaluative criteria can transform the methodologist's view of theory-appraisal. The power of Holton's work derived largely from his insight that science was in some ways similar to the arts: but this power was somewhat dissipated when he admits that in his view science is similar to the arts only in the process of creation, and not in the evaluation of the products. It is partly in the extension of aesthetic categories to the logical positivist's 'context of justification' that the present treatment claims originality over Holton's dynamic of themata.

The august confines of positivism are transcended in the study of the

\textsuperscript{21} Holton (1973), pp. 21-9. For further discussion of Holton's thematic analysis see e.g. Losee (1987), pp. 135-41.
evvaluational role of aesthetic criteria and of their consequent justificatory power in scientific practice.

4. The disjunction of aesthetic and epistemic values

Heisenberg on one occasion aired the possibility that there might exist what he termed ‘aesthetic criteria of truth’. The passage in which this suggestion is contained is the following:

If nature leads us to mathematical forms of great simplicity and beauty – by forms I am referring to coherent systems of hypotheses, axioms, etc. – to forms that no one has previously encountered, we cannot help thinking that they are ‘true,’ that they reveal a genuine feature of nature. [...] You may object that by speaking of simplicity and beauty I am introducing aesthetic criteria of truth, and I frankly admit that I am strongly attracted by the simplicity and beauty of the mathematical schemes with which nature presents us. You must have felt this, too.22

Although this suggestion has already received an implicit response above, it is so distinctive and so contrary to the thrust of the present treatment that it is advisable to consider it explicitly. In this section the suggestion concisely formulated by Heisenberg will be expanded and elucidated, and three reactions to it will be provided.

By ‘aesthetic criteria of truth’ Heisenberg presumably intended criteria of theory-evaluation which exhibit two characteristics: a) they are properly applied to or operate upon features of theories of a bona fide aesthetic or perceptual nature, rather than to logico-empirical features of theories; b) their evaluative application to a theory yields a judgement or estimate of its likely truth or verisimilitude. One may broaden the suggestion of Heisenberg to envisage that such putative criteria are applied to gauge not exclusively the likely truth of theories but also as an alternative the degree of the theories’ likely future observational success: this broadening accommodates the views of those, like

22 Heisenberg (1971), pp. 68-9. This passage is contained in the transcription of a rejoinder which Heisenberg made in conversation to Einstein in 1925. The idea that there may exist aesthetic criteria of truth will be referred to in this section as ‘Heisenberg’s suggestion’, though it is not implied here that Heisenberg propounded this hypothesis deliberately or on occasions other than the one on which he spoke to Einstein in the terms here reported.
antirealists, who believe that it is illegitimate to predicate truth or truthlikeness of theories.

Heisenberg’s suggestion would be borne out if there existed aesthetic or perceptual features of theories such that their possession by a given theory were indicative of that theory’s observational success. If – following the discussion of chapter 2 above – one considers that there exists a number of logico-empirical features of theories which are constitutive of observational success, Heisenberg’s suggestion reduces to the proposition that there exist aesthetic features of which the possession by a theory is correlated with that theory’s possession of the logico-empirical features constitutive of observational success. In short, according to Heisenberg’s suggestion, a theory’s possession of certain aesthetic features would make it more likely that the theory should in future be revealed observationally successful.23

Such a correlation undoubtedly holds, for instance, in the answers to questions in university examinations in the natural sciences: the appearance in an attempted answer to an exam question of aesthetically displeasing features such as irreducible mathematical expressions or equations which admit no analytical solution is most often an indication that the attempted solution is ill-conceived, or at least harbours a mathematical error some way upstream of the formula reached. Similarly the beauty or ‘neatness’ of the emerging configuration of some of the answers essayed by the candidates justifiably persuades them that they are on the right track.24 Needless to say, the fact that the attempted solution’s possession of such aesthetic features can reliably be taken as an indication of its accuracy is a consequence of the artifical context in which the question has been framed, and possibly of the predispositions of the examiner for simple and easily-graded solutions; problems in open-ended scientific research are clearly not amenable to the same considerations which one would turn upon examination.

23 The possibility that a theory’s possession of given aesthetic features is somehow correlated with its proximity to the truth is implicitly advanced also in A.M. Taylor (1966), p. 38: ‘The elegant beauty of the theoretical edifice [of general relativity] is thought sufficient reason for believing it to be true.’ Watson (1968, p. 210) speaks similarly of a theory’s being ‘too pretty not to be true.’ Perhaps the most forthright formulation of Heisenberg’s suggestion has come in the headline to Trefil (1986): ‘The Most Beautiful Theories Are the Truest’. This work is a review of Holton (1986); its title appears to bear little relevance to the content of either Trefil’s article or Holton’s book, and was doubtless the creation of an editor under the spell of Keats.

24 Many university textbooks in the physical sciences explicitly exhort students to awareness for the elegance of the solutions which they propose to examination questions: Pippard (1972, pp. 10-1) teaches his readers two ways – one ‘straightforward’ and the other ‘elegant’ – by which to reach a solution of a typical examination problem in kinematics.
CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

questions. The issue here is whether the use of aesthetic features as indicators of truthlikeness which is successfully made by examination candidates is equally appropriate to the methodology of the research scientist.

The present treatment confronts Heisenberg's suggestion with three distinct responses. The first refers to the structure of the investigation pursued in the present treatment; the second demands that those who endorse Heisenberg's suggestion should elucidate a view of aesthetic features of theories which would enable the suggestion to be realized in scientific practice; the third brings to bear historical evidence in an examination of the suggestion's plausibility.

Consider first the manner in which the above treatment discerned the existence of a conceptual vacancy opened by the project of adequately reconstructing the history of science. Chapter 2 recounted how the community of historians and philosophers of science have come to acknowledge the hopelessness of attempting to reconstruct the past record of science by appeal to logico-empirical criteria of theory-evaluation alone. In this acknowledgement the community recognized that, in order to render justice to the history of science, reconstructions are compelled to appeal to a further class of criteria of theory-assessment. Criteria of this class have to differ from the previously-identified logico-empirical criteria: after all, constant recourse has been made in the reconstruction of the history of science to the logico-empirical class of criteria, in a variety of differing formulations, and that class has been found insufficient. Something new is required in the project of reconstructing the history of science.

Under the impulse of this apprehension the above treatment turned to consider aesthetic criteria of theory-evaluation. Chapter 4 was able to discern in the history of philosophical aesthetics a tradition which interpreted aesthetic perception as characterized by lack of concern for utilitarian features of the objects of that perception; but independently of and prior to the historiographic discovery of that aesthetic tradition, the search which had been undertaken for a further class of criteria of theory-evaluation had stipulated that the criteria which constituted the object of the search were non-empirical.

It is at this stage in the argument that one considers what has above been called Heisenberg's suggestion, that there might exist aesthetic criteria of theory-evaluation which function as indicators of a theory's empirical virtues. Clearly the first response which the present treatment must move to such a suggestion is to point out the distance which separates it from its own concerns. Heisenberg's suggestion is at root the postulation of an association between truth and beauty, which – if it is to support rigorous analysis – must allude to a mechanism by

73
which the association is assured. Such a mechanism could be adduced most plausibly by demonstrating how a theory’s endowment of logico-empirical virtues ensures its simultaneous possession of certain aesthetic features. But then a theory’s lack of specified aesthetic features would be indicative of its lack of logico-empirical virtues: so any explanation of the mechanism by which beauty and truth were associated would conflate the two categories of logico-empirical and aesthetic evaluative criteria, since in virtue of the mechanism one could utilize aesthetic criteria to determine the logico-empirical worth of the theory under assessment. Since the central thrust of the present treatment is to offer aesthetic criteria of theory-evaluation in fulfilment of the search for a class of evaluative criteria separate from the previously-established logico-empirical criteria, it can hardly be expected that Heisenberg’s suggestion will be susceptible of organic incorporation into the structure of this treatment. Heisenberg’s suggestion is not merely of an opinion contrary to the ones propounded by the present treatment: it is, more radically, foreign to the manner in which the problem tackled by the present treatment has been constituted. Someone desirous of embracing a view of the incidence of aesthetic considerations in scientific practice may of course choose freely between Heisenberg’s suggestion and the ones advanced here; but the choice between those alternatives must be taken before following the present treatment to the point which the reader has now attained. From the perspective of the present treatment, thus, Heisenberg’s suggestion is not a contribution to the debate in hand but rather a denial of the terms constituting the current debate and a – fully legitimate – attempt to undertake a wholly different conceptual itinerary which happens to share with the one pursued here only a concern for features of theories termed ‘aesthetic’.

If the relation between Heisenberg’s suggestion and the present treatment is understood in these terms, the present treatment’s second response to that suggestion arises naturally. It is legitimate to demand of those who embrace Heisenberg’s suggestion that they adduce a mechanism by which the copresence of logico-empirical and of aesthetic features in theories is assured, and in virtue of which it is possible to envisage the existence of ‘aesthetic criteria of truth’. It is hardly sufficient for Heisenberg to intimate that given ‘mathematical forms of great [...] beauty’, ‘we cannot help thinking that they are “true”’: to incorporate his suggestion into a characterization of scientific rationality, we must be told what is the warrant of this belief, what it is about the beauty of a mathematical form which makes it truer than less beautiful ones. Such a mechanism has not so far been proposed by those few writers who have followed Heisenberg in this
CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

matter; until the details are thus clarified, it is difficult to begin to evaluate the merit of Heisenberg's suggestion.

The third of the responses to Heisenberg's suggestion which were promised above brings historiographic data to bear upon the issue. If a model of the practice of theory-evaluation which incorporated Heisenberg's suggestion were capable of accurately reconstructing substantial spans of the history of science, one would expect to find in the historical record instances in which scientists referred to a theory's possession of certain aesthetic features as evidence for supposing that the theory would demonstrate logico-empirical validity. This expectation is not fulfilled by historiographic findings. On the contrary, the historical record indicates that scientists generally consider their aesthetic judgements of a theory to be autonomous from their empirical verdicts on it: in general, their aesthetic and empirical evaluations of a given theory differ. This suggests that the intuitions of scientists hold aesthetic criteria to constitute a set of evaluative standards altogether disjoint from that of empirical criteria.

For example, in writing that 'Einstein's theory [of general relativity] has the very highest degree of aesthetic merit: every lover of the beautiful must wish it to be true', H.A. Lorentz implicitly acknowledges that the theory of relativity may be beautiful and yet not true, that it may satisfy one of the classes of evaluative requirement but not the other. In this belief he is joined by E. Rutherford who in 1932 wrote 'The theory of relativity by Einstein, quite apart from any question of its validity, cannot but be regarded as a magnificent work of art', thus endorsing the notion that the aesthetic virtues of Einstein's theory were a matter separate from that of its empirical validity.

One of the most celebrated instances of such a evaluative disjunction between, to put it briefly, truth and beauty appeared in physicists' reaction of a theory in quantum mechanics developed in the 1940s. Quantum electrodynamics numbers among the most successful modern accounts of a body of experimental data: it gives values for such physical quantities as the Lamb shift and the anomalous magnetic moment of the electron which agree with measurement within the bounds of experimental accuracy, which is of a few parts per million. However, quantum electrodynamics yields these brilliant predictive results only after certain infinities, which appear during the calculations in the values for the electron mass and charge, are excised by a mathematical procedure developed by

---

26 Quoted in D. Wilson (1983), p. 432; for further details of the context of this remark see ibid., p. 594.
J. Schwinger, R.P. Feynman, F.J. Dyson and others and named ‘finite renormalization’. Because of the need for this procedure, quantum electrodynamics struck and continues to strike many physicists as aesthetically displeasing, especially compared to putative finite field theories which would be free of such blemishes. Dirac, who had himself published outstanding work in quantum field theory especially in the 1920s, persistently refused to accept renormalized quantum electrodynamics as a satisfactory physical theory. He manifested no reservation about its ability to account for experimental data; his scepticism was concerned entirely with what he considered to be the unacceptable inelegance of the doctoring which is necessary to draw from it determinate predictions. He expressly noted the cohabitation in quantum electrodynamics of empirical success and aesthetic blemish, and directed his repeated criticism exclusively to the latter feature of the theory while remaining ready to pay tribute to the purely empirical success of the theory’s predictions. He wrote in a formulation typical of his concerns: ‘Recent work by Lamb, Schwinger, Feynman and others has been very successful [...] but the resulting theory is an ugly and incomplete one, and cannot be considered as a satisfactory solution of the problem of the electron.’

There exists also a class of historical instances in which theories judged aesthetically pleasing have simultaneously been recognized as empirically inadequate. Schrödinger has aesthetic praise but empirical criticism for the Lamarckian theory of evolution: he writes that it is ‘beautiful, alluring, encouraging and invigorating’, but adds, ‘Unhappily Lamarckism is untenable. The fundamental assumption on which it rests, namely, that acquired properties can be inherited, is wrong.’ The same disjunction is asserted more recently by D.

27 An account of the development of quantum electrodynamics, of the invention of the procedure of renormalization and of its empirical successes is given by Weinberg (1977), pp. 21-30.

28 Dirac (1951), p. 291. Shanmugadhasan (1987, p. 53) adduces a similar instance of Dirac’s aesthetic repudiation of this theory: he writes that in 1945 Dirac ‘emphasized that he did not believe his quantum electrodynamics was the right theory because it was so complicated and ugly.’ The role played by Dirac vis-à-vis quantum electrodynamics – the role of an esteemed senior physicist who acknowledges the empirical power of a new theoretical development but cannot bring himself to accepting its aesthetic features – was played at around the same time by Einstein in regard to quantum physics as a whole: while Einstein recognized the empirical success of the new quantum theory he found its abandonment of determinism irreducibly ugly. Chapter 8 below, which presents a general theory of scientific revolutions with regard to the aesthetic commitments of factions of the scientific community, will suggest that this position is frequently and – to some extent – predictably taken by some scientists at times of revolutionary crisis. Chapter 10 will outline Einstein’s reactions to quantum mechanics at greater length.

Sciama in commenting on the cosmological steady-state theory of F. Hoyle and others: 'It is very beautiful but it is now in serious conflict with observation.'\(^30\) J. Rosen generalizes the observation that a theory with less aesthetic appeal may exhibit superior empirical virtues:

> If we eavesdrop on private discussions among scientists, we might hear expressions such as, 'This is a beautiful theory (of ours)!’ or, 'His theory is rather ugly.' Both theories might be equally good, in that they both explain the same natural phenomena equally well. In fact, the 'ugly' theory might even be better.\(^31\)

Each such instance lends weight to the proposition, embodied throughout the present treatment, that scientists hold the aesthetic evaluation of theories to be an act separate from any empirical judgements of them. The correlation between elegance and empirical validity which holds so perceptibly in the realm of solutions to exam questions does not carry over to scientific theories in general. This observation plays an important role in erecting aesthetic evaluations as an object of inquiry on the part of philosophy of science autonomous of the study of logico-empirical criteria.

5. Aesthetic criteria and the underdetermination of theory-choice

There exists a conceptual route of considerable interest, different from those explored thus far, by which some methodologists of science have acquired a commitment to the beliefs that aesthetic criteria of theory-evaluation operate in theory-choice and that their verdicts are independent of those handed down by logico-empirical evaluative criteria. This route originates in the recognition of the problems which would arise for philosophy of science – especially for realist doctrines of science – from instances of theory-choice underdetermined by empirical data. Logico-empirical evaluative criteria would be insufficient to

---

\(^30\) Quoted in Osborne (1986), p. 12. A similar phrase is attributed to Sciama in Kippenhahn (1984), p. 153: 'The steady-state theory has a sweep and beauty that for some unaccountable reason the architect of the universe appears to have overlooked.' The import of these statements is identical: the steady-state theory's failure to be instantiated in the universe is revealed by a predictive inadequacy which is quite independent of its aesthetic virtue. For bibliographic help on this topic I am indebted to Dr D.W. Dewhirst of the Institute of Astronomy, University of Cambridge.

\(^31\) Rosen (1975), pp. 120-1.
adjudicate among two or more competing theories which exhibited the same predictive power over all possible evidence but possessed some incompatibility (e.g. radically different ontological commitments) which prevented their being considered alternative expressions of a common theoretical substructure. If this underdetermination were not overcome, it would presumably be impossible to maintain the realist view of scientific progress as leading to a uniquely-determined ‘truth of the matter’ about the natural world.\textsuperscript{32}

If logico-empirical evaluative criteria are feared to be insufficiently discriminating to determine theory-choice in such cases, that unpalatable outcome may yet be averted by making recourse to a set of evaluative criteria separate from and additional to the logico-empirical criteria, in the expectation that the joint application of the two sets of evaluative criteria will prove sufficiently discriminating to determine uniquely any case of theory-choice which is refractory to logico-empirical criteria alone.\textsuperscript{33} Those philosophers who have pursued this – the intuitively most obvious attempt to avert the underdetermination of theory-choice – have often turned to aesthetic or quasi-aesthetic criteria of choice from which to construct the second or supplementary evaluative canon, even if generally such utilitarian recourse to aesthetic criteria has not been underpinned by a pondered view of the incidence of aesthetic considerations in wider scientific methodology. On this prescription, in the event where rival theories appear equally truthlike one should like Paris choose the most beautiful for somehow that would hold the most promise. F. Rohrlich endorses this course of action:

There is [...] great beauty in a physical theory. [...] It is that beauty which affects the credibility of one theory over another in the absence of more stringent criteria. For instance, the general theory of relativity is so beautiful that it is preferred over rival theories as long as those rival theories cannot account any better for the empirical facts.\textsuperscript{34}

R. Swinburne entrusts this role of tie-breaker to a criterion of simplicity of which the content may be determined partly by aesthetic predispositions.\textsuperscript{35} Though moving from a philosophical perspective very different from that of Swinburne, P.K. Feyerabend shares with him both the identification of certain cases of theory-

\textsuperscript{32} For a fuller discussion of underdetermination and the dangers which it poses to realism see Newton-Smith (1978).

\textsuperscript{33} Newton-Smith examines and appraises this strategy for overcoming the underdetermination of theory-choice \textit{ibid.}, pp. 76-7.

\textsuperscript{34} Rohrlich (1987), pp. 13-4.

\textsuperscript{35} Swinburne (1968), p. 21. In the next chapter ‘form of simplicity’ will be identified as a possible component of an aesthetic canon of theory-assessment.
choice as problematic and the apprehension that in such cases it is natural to turn to aesthetic and similar judgements to break the *impasse*. Feyerabend describes the difficulty of choosing between incommensurable theories as follows:

None of the methods which Carnap, Hempel, Nagel, Popper or even Lakatos want to use for rationalizing scientific change can be applied, and the one that can be applied, refutation, is greatly reduced in strength. What remains are aesthetic judgements, judgements of taste, metaphysical prejudices, religious desires, in short, *what remains are our subjective wishes*.  

A later chapter will show that Feyerabend errs in considering aesthetic judgements as no more than 'our subjective wishes', and more broadly in failing to investigate mechanisms for the construction, entrenchment and operation of aesthetic criteria of theory-assessment, which would greatly assist the project of 'rationalizing scientific change'. Nonetheless it is encouraging to discover that some philosophers of science have reserved attention for aesthetic categories, however tangential their interest in the matter.

6. The place of aesthetics in Einstein's view of theory-assessment

Einstein is the author of numerous celebrated adages on the practice of theory-formulation and assessment. Many refer crucially to aesthetic requirements of theories, and the more univocal of these pronouncements are cited elsewhere in the present treatment.

Alongside all such occasional pronouncements Einstein left a sophisticated and complex view of the procedures of theory-evaluation, contained mainly in his 'Autobiographical Notes'. The degree of coherence between his frequent recourse to aesthetic criteria in practice and his more detached perception of theory-assessment is not clear, for whereas he on occasion quite unambiguously judged theories on aesthetic grounds he appears at first sight not to have dedicated an

---

36 Feyerabend (1975), pp. 284-5; emphasis in the original.
37 Holton, who has long studied Einstein's methodology, identifies in his practice of theory-formulation and assessment concern for the following set of theoretical *desiderata*, many of which have an aesthetic or quasi-aesthetic nature (1981, p. 15): 'Primacy of formal (rather than materialistic or mechanistic) explanation; unity or unification; cosmological scale in the applicability of laws; logical parsimony and necessity; symmetry (as long as possible); simplicity; causality (in essentially the Newtonian sense); completeness and exhaustiveness; continuum; and of course constancy and invariance.'
overt and specific niche to those grounds in his more systematic statements on the topic. The methodological portrait of Einstein which implicitly emerges from the present treatment would therefore be incomplete without a brief review of those statements, which will be followed by an indication of the implicit locus within them of aesthetic criteria and a remark on the affinities between Einstein's view of theory-assessment and the present treatment.

In Einstein's view the corpus of science consists of a structure of interrelated concepts removed from sense-experience. These concepts are formulated by conjecture constrained by experimental data, to which they are related by logical implication. A scientific theory thus exhibits two classes of relation: the set of relations between the invented concepts, and the set of relations between those concepts and items of sense-experience. Corresponding to the two sets of relations are the two levels on which theory-assessment may be conducted: an internal level evaluating the theory's inherent conceptual structure, and an external one pertaining to the relationship of the theory to experiment.38

Einstein summarizes the requirements of the external appraisal by the remark that 'the theory must not contradict empirical facts.'39 There are however according to him two reasons for the insufficiency of concurrence with experiment as an evaluative criterion, and simultaneously for the need of an internal level of theory-assessment. First, it is always possible to eliminate discrepancies between theory and experiment by means of ad hoc hypotheses, which, though at times admirably satisfying the requirements of the external level of assessment, reduce the overall scientific worth of the theory.40 Secondly, in modern physics the chains of inference from the principles of a theory to its observational consequences become increasingly long and complicated, so that 'the confrontation of the implications of theory by the facts becomes constantly more difficult and more drawn out.'41 Einstein's perception that theory-evaluation on the external level is becoming more arduous adds a pragmatic reason for the use of the internal level, alongside the philosophical reason that the external level may on its own underdetermine theory-choice.

The internal level of theory-assessment 'is not concerned with the relation to the material of observation.'42 Einstein mentions two criteria which operate at
the internal level of theory-assessment: the ‘naturalness’ or logical simplicity of a theory, and the definiteness of its descriptions.

The simplest conceptual system is the one containing the least number of arbitrarily chosen elements. Newtonian mechanics contained as fundamental statements both principles constituent of Euclidean space and laws of force.\textsuperscript{43} The primary force, gravitation, was described by an inverse-square distance law, chosen by Newton from all possible such statements; this choice was arbitrary since it was not imposed or suggested by the other elements of the system. Einstein argued that this degree of arbitrariness could be reduced by reformulating the theory. The least-power spherically symmetric solution of Laplace’s equation for a potential is an inverse first-power function of distance. Differentiation of this potential function yields an inverse-square distance law of force. This law has now come to be determined by the geometry of Euclidean space: the connection reduces the arbitrariness of the theory.

While this example is the one adduced by Einstein to illustrate the application of a simplicity-criterion to theory-evaluation, a more incisive instance may be culled from Einstein’s own work. Newtonian theory possessed two operational definitions of mass, inertial and gravitational, logically independent of one another; Einstein’s general relativity posits the identity of the two qualities, reducing by one the number of arbitrary variables in the description of a general mechanical system.

The second criterion of internal theory-assessment is the degree of definiteness of the theory’s assertions: Einstein prescribes that among theories of equal simplicity one should prefer that which makes the most definite claims about the systems described.\textsuperscript{44} Einstein offers no example of a theory embodying some degree or other of this virtue, but P. Barker suggests that he may have had in mind the greater definiteness of the field-formulation of Newtonian mechanics respect to its formulation in terms of inter-particle forces.\textsuperscript{45} The latter formulation leaves unspecified the states of a system at instants of particle collisions, an indeterminacy not present in the former version.

Einstein’s internal criteria of theory-evaluation are – as he admits – not so precise as to permit a quantitative comparison between the conceptual parameters of rival theories: Einstein particularly dismisses numerical measures of logical simplicity because they require an arbitrary decision on what counts as a logically

\textsuperscript{43} Einstein presents this example \textit{ibid.}, pp. 29-33.
\textsuperscript{44} \textit{Ibid.}, p. 23.
CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

independent expression. More modestly Einstein describes his internal criteria as involving 'a kind of reciprocal weighing of incommensurable qualities.'

As already mentioned, the above speculation on theory-assessment in the 'Autobiographical Notes' reserves no specific attention for aesthetic evaluative criteria. In the light of the evidence of his concerns presented in section 1 of this chapter and elsewhere, it would be hazardous to conclude from this apparent omission from the 'Notes' that Einstein attributed to aesthetic factors an inessential weight, or perhaps a role only in theory-formation rather than assessment. The interpretation of the 'Notes' which accords best with Einstein's other pronouncements is obtained by reading within them an implicit description of the role of aesthetic evaluative criteria. If the external level of theory-assessment is empirical, it is reasonable to conclude that Einstein located aesthetic concerns in the internal level, which pertains to the logical interrelation of the concepts of theories. This interpretation coheres with Dirac's view that Einstein referred to the twin evaluative categories of 'agreement with observation' and 'beauty in the mathematical foundation', as well as with Bergmann's ascription of Einstein's adoption of a theory to 'its confirmation by experiment and observation' and 'the aesthetic appeal of the finished theory.' If this reading is followed, Einstein's talk in the 'Notes' of the 'naturalness' of a theory assumes at root aesthetic connotations, a plausible interpretation.

While this interpretation of the 'Notes' helps to clarify Einstein's methodological dispositions, the present treatment does not seek Einstein's blessing with its theses: the view of aesthetic factors developed here is wholly separate from that of Einstein. The present view shares with Einstein's the partition of criteria of theory-assessment into two classes, one but only one of which pertains to the empirical performance of theories. The characterization of the other class differs in the two treatments, and in this treatment includes aesthetic factors of a range wider than Einstein probably envisaged. Most important, Einstein leaves uninvestigated both the genesis of criteria of internal evaluation and the case of conflict between the two classes of criteria in theory-choice: these topics are the subject of later chapters.

46 Einstein (1949), p. 23. Barker twice (1981, pp. 139, 141) erroneously transcribes this expression as a 'reciprocal weighing of incommensurable quantities', fuelling the confusing impression that Einstein perhaps envisaged some numerical measure of theoretical virtues after all.

CHAPTER FIVE: THEORY-EVALUATION ON AESTHETIC GROUNDS

7. Alternative justificatory roles of aesthetic features

The previous sections suggested that scientific communities adduce a theory’s possession of certain aesthetic qualities among the grounds for the adoption of that theory. A different use of aesthetic features to justificatory ends in scientific practice may however be envisaged. Here three alternative ways in which justificatory appeal may be made to aesthetic features will be outlined; although consideration of these justificatory powers of aesthetic features in science is not relevant to the main present theses, and therefore will not be pursued at great length, it is valuable to compare the approach of the present treatment with certain other views.

N. Jardine suggests that there may exist two forms of appraisal in scientific practice in which reference is made to aesthetic features in order to assist in the choice between competing theories, but where the aesthetic features to which reference is thus made are qualities not of the theories under evaluation but rather of other entities connected to them, as will be explained now. The first of the two types of aesthetic appraisal envisioned by Jardine attributes aesthetic value to phenomena viewed in the light of an explanatory theory, in a procedure reminiscent of the perception of aesthetic value of the second degree of abstraction discussed by Hutcheson and described in chapter 3 above. The act of evaluation which makes use of this device attributes preference to the theory under assessment to the degree to which it is able to ‘bring out’ aesthetic value in the phenomena which constitute its subject-matter. For an example of this form of appraisal Jardine refers to the final paragraph of Darwin’s *The Origin of Species*. The intricate but organic structure of the ‘entangled bank’ is by Darwin attributed aesthetic value: in Jardine’s view this value – while being a feature of the phenomena which constitute the subject-matter of the theory of adaptation and natural selection rather than of the theory itself – militates in favour of the adoption of that theory as one well able to ‘bring out’ those features of the object.

The second form of aesthetic appraisal in science envisaged by Jardine confers aesthetic value to theory-based representations of phenomena or experimental data, such as graphs and diagrams. Each of the competing theories

---

49 Further on Darwin’s perception of aesthetic images in his subject-matter see e.g. Gruber (1978).
50 The aesthetic principles involved in the design and perception of such science-based representations of phenomena are discussed in Tufte (1983), pp. 177-90.
of a sector of reality will generate representations of the phenomena within its scope which can be aesthetically evaluated as objects of perception in their own right. Jardine suggests that the comparative assessments of these representations will reflect upon the theories which spawned them: this form of appraisal recommends preference to be given to those theories which produce representations of phenomena of the greatest aesthetic appeal.

It will be noted that although the outcome of acts of evaluation of these two forms is an alteration of the degree of preference accorded to theories in instances of theory-choice, the objects of the aesthetic judgements passed in these acts are entities other than the theories themselves: the objects of the evaluation are natural phenomena in evaluations of the first kind and graphic representations of the phenomena in the second. The theory which reconstructs the incidence of aesthetic considerations in science as the operation of appraisals of these two kinds thus does not venture beyond the notion of discerning aesthetic value in concrete objects, such as natural scenes or graphic representations of phenomena. The perception of beauty in nature or in concrete portrayals of nature is familiar to the traditional conception of aesthetics as a discipline constituted around the problems of art criticism. The present treatment has on the contrary suggested that the extension of the application of aesthetic categories to abstract objects or intellectual constructs, including scientific theories, might further assist the enterprise of modelling scientific practice.

The third possible justificatory role of aesthetic features in scientific practice to be mentioned in this section is even further removed from the methodological concerns of the present treatment. It has been suggested by some that scientists see in the aesthetic values of their theories a force justifying not the embrace of one theory to the neglect of another, but the pursuit of science: the aesthetic gratification caused by the contemplation of a scientific theory is in that interpretation deemed not grounds to prefer that theory to others, but a reason to engage in scientific research as a whole, as a life’s-work in the course of which a dose of aesthetic gratification will be obtained. This is a form of transcendental justification, in that features of the products of the enterprise come to justify its pursuit. This view of the justificatory valence of aesthetic features of theories is entertained by R. Penrose who enquires after the ‘justification for doing mathematical research’ and concludes that ‘basically, the motivations turn out often to be aesthetic ones; so often a subject is pursued simply for the pleasure
that it gives. This motivational use of aesthetic features of theories, which in any case is compatible with their use in theory-assessment, departs excessively from the concerns of the present treatment and will not be considered further.

---

52 The role of aesthetic delight in scientists' motivations for pursuing research is mentioned in Chandrasekhar (1987), passim.
Chapter Six
AESTHETIC FEATURES
TO HAVE WON COMMENDATION

1. The specificity of descriptions of aesthetic canons

At what level of specificity must an aesthetic canon be described before its effect in theory-assessment can be gauged? When the topic of aesthetic evaluations of scientific theories is broached in general conversation, the response commonly obtained is, 'Oh, you mean scientists' preference for theories which are symmetric and simple?'. Let us leave aside for a moment the issue whether symmetry and simplicity are indeed aesthetic properties of theories, and ask instead the question: would a description of an aesthetic canon as consisting of preferences for theories which exhibit symmetry and simplicity permit a study of the effect of applying that canon in cases of theory-evaluation?

One must answer this question 'no', on the grounds that the description of an aesthetic canon which it proposes is insufficiently determinate for that purpose. Any or almost any theory could be said to possess 'symmetry' or 'simplicity', so the information that a canon consists of preferences for theories which exhibit these properties would be insufficient, for instance, to predict the verdict of the canon in a case of theory-evaluation. A more specific description of the content of the canon is required for that purpose: one needs to be told which form or forms of symmetry or simplicity are held by the canon to be especially desirable in theories.

The demand for specificity is made whenever we attempt a description of an aesthetic evaluative canon. We are told for instance by E.M. Forster that the aesthetic canon by which he judges novels attributes value to symmetries in the plot.¹ This information does not shed much light on Forster's actual preferences: as far as we have been told, his canon might prompt him to extend preference to the novels of Hardy because in them the landscape has a 'mood' which appears

¹ Forster (1927), pp. 136-7.
to mirror the mood of the protagonists, or those of Dickens because in them
events are recounted from the point of view of more than one protagonist. Before
we can study the aesthetic evaluations which Forster gives of novels, we must
know which form of the property ‘plot-symmetry’ his canon specifies as
particularly desirable in novels. Foster implicitly provides this information in
commending *The Ambassadors* by Henry James because in it two of the
protagonists exchange places, so that the novel assumes ‘the shape of an hour­
glass’.\(^2\)

Just as unqualified talk of symmetry is insufficiently specific to determine
the preferences prompted by the aesthetic evaluative canon of the literary critic,
it is insufficient to specify those of the scientist: the form of the symmetry sought
by the aesthetic canon of a community at a given time must be stated before one
can know the prescriptive content of that canon.

In the study of the composition of aesthetic canons of theory-assessment
which will occupy the greater part of this chapter, entities like a certain form of
symmetry will be termed ‘aesthetic features’ of theories and entities like
symmetry will be termed ‘aesthetic feature-classes’. From what has already been
said it follows that an aesthetic evaluation of a theory will be based on that
theory’s possession or failure to possess certain aesthetic features, not certain
feature-classes. Two aesthetic evaluative canons which prescribe different forms of
symmetry to be desirable in theories will be said to differ in attaching value to
different aesthetic features, but to resemble one another in referring to the same
feature-class.

The relation between a given feature-class and a feature which belongs to it
is similar to that in political theory between the notion of e.g. ‘freedom’ on the
one hand and the liberalistic or Marxist notions of freedom on the other. The
proliferation of notions of ‘freedom’ to which weight has been attached in
political thought through the ages resembles the proliferation of forms of
symmetry to which value has been attributed in the appraisal of scientific
theories. To know that someone’s ideology sets great store by freedom does not
allow the policy-choices which that agent will make to be predicted: similarly, the
knowledge that symmetry carries great weight in a scientific community does not
reveal the theory-choices to which that community will be prompted. In each case
knowledge is required of the precise form of the notion to which preferential
weight is attributed.

\(^2\) Ibid., p. 137.
The previous chapter established the hypothesis that scientists' appraisals of and choices between theories have been affected by those theories' possession of certain aesthetic features. The remainder of this chapter will enumerate and discuss some of the aesthetic features of theories which have thus influenced theory-choice in the history of science. The aesthetic features cited will be categorized into four feature-classes: simplicity, symmetry, analogical tractability, and consonance with metaphysical presuppositions. It is claimed neither that these exhaust the set of feature-classes to which the aesthetic evaluation of theories has made recourse in the history of science, nor that the features to which scientists have referred in certain cases of theory-choice can appropriately be allocated among the four feature-classes in only one way.3

This chapter will furthermore advance and support three claims about the features to which it will refer:

(i) that these features are aesthetic in nature in the sense required by chapters 4 and 5, i.e. that their possession by theories is uncorrelated with those theories' possession of logico-empirical virtues;

(ii) that scientists in history have on many occasions evaluated theories on their possession of aesthetic features, and that many of these evaluations contributed to determine the direction in which the corpus of science was to evolve;

(iii) that scientists' aesthetic tastes have shown great temporal mutability, or more precisely that there have been considerable changes over time of the set of aesthetic features which it has been held desirable that theories should possess.

A few remarks may help to explain why these claims are thought worthy objects of advocacy. The list of features of theories contained in the next four sections was compiled from historiographic sources: a survey of the history of science threw up the items on this list as features of theories of prima facie aesthetic or non-empirical nature on which theories were on occasion assessed. But that the theoretical features upon which attention was in this way drawn are 'aesthetic features of theories' in the sense presumed in chapters 4 and 5 above requires demonstration. To argue for the truth of claim (i) is to attempt to show

3 In particular the discussion among Ptolemaic, Copernican and Keplerian astronomers about the degree of harmony afforded by the postulation that celestial bodies traced circular paths may be seen as debates about the feature-classes of simplicity or of the symmetry of astronomical theories or about the feature-class of 'coherence with metaphysical presuppositions.'
that the features suggested by the historiographic survey are indeed relevant to
the study of the practice of aesthetic theory-evaluation upon which the present
treatment is embarked. To argue for the truth of claim (iii), on the other hand, is
to contribute to demonstrating that the practice of aesthetic theory-evaluation is
an influential component of scientific methodology, that an account of the history
of science which failed to refer to scientists' aesthetic evaluations of theories
would be considered incomplete, and hence that this practice demands study by
the historian and philosopher of science. Finally, that scientists' aesthetic
evaluative canons have exhibited great mutability over time is the observation
that will constitute one of the explananda of the next chapter; to argue for the
truth of claim (iii) is thus to strive to establish the occurrence of some of the
historical phenomena to be discussed in chapter 7.

3. Simplicity

Aspects of the property of simplicity exhibited by scientific theories have been the
subject of innumerable treatments. Most of these endorse one of two views. The
first is that in comparative theory-evaluation the degree of simplicity of a theory
is indicative of its proximity to the truth: a simpler theory is ceteris paribus more
likely to be true or empirically adequate than a less simple one. On this view the
degree of simplicity of a theory is a logico-empirical feature of the theory, and it
is on the basis of a theory-evaluation conducted on logico-empirical grounds that
one chooses ceteris paribus to adopt the simplest theory among several that may
be proposed. The present treatment records here its endorsement of this view on
the grounds of the arguments outlined in the treatment of the logico-empirical
criterion of simplicity contained in chapter 2.4

The second view about simplicity which is commonly endorsed is that the
form of simplicity of a theory is an aesthetic property of it, or at least a property
appeal to which is of no use in attempting to establish the likely degree of future

4 Chapter 2 contained also examples drawn from the history of science of the appeal
to the simplicity of a theory as evidence of its empirical adequacy.
CHAPTER SIX: AESTHETIC FEATURES TO HAVE WON COMMENDATION

observational success of that theory. On this view the choice to adopt one theory in preference to another in virtue of the forms of simplicity which they exhibit is made on criteria of aesthetic taste.

Many writers appear to believe that the two views cited here are incompatible, and more particularly that, if one accepts the degree of simplicity of a theory to be a truth-related, logico-empirical property of that theory, any simultaneous attempt to submit a theory to aesthetic evaluation on account of its simplicity is illegitimate. The attempt to establish that this belief is mistaken, and that it is legitimate to attribute to simplicity both logico-empirical and aesthetic aspects, will engage the remainder of the present section.

The argument proceeds initially per absurdum. Those who believe that the simplicity of a theory is indicative of the empirical virtue of that theory believe also – correctly – that it is the degree of that theory’s simplicity which is correlated to its proximity to the truth or to its degree of empirical adequacy. Under this presupposition, for it to be illegitimate to ascribe to the simplicity of a theory any aspects not correlated with the logico-empirical evaluation of the theory, a specification of only the degree of simplicity exhibited by that theory would have to constitute a complete description of the simplicity of that theory. In other words, for it to be illegitimate to ascribe to the simplicity of a theory any non-empirical aspects, the specification of the value of one parameter only – the degree of the simplicity exhibited by a theory – would have to suffice to obtain an exhaustive specification of the simplicity by which the theory is possessed.

But findings of the history of science suggest that to specify completely the simplicity by which a theory is possessed requires much more information than this. As evidence for the proposition that the specification of the degree of the simplicity of a theory would fail on its own to constitute an exhaustive characterization of the simplicity of that theory, consider the following historical instances of the appeal to simplicity in theory-evaluation.

Copernicus wished astronomical theory to be simple in appealing to only

---

5 Einstein for instance appeared to believe that ‘simplicity was equivalent to beauty’, in the words of Elkana (1982, p. 222); E.O. Wilson (1978, p. 11) considers the simplicity of a generalization to be an important component of its elegance or beauty. The absence of a correlation between the simplicity of a theory and its truthlikeness has been affirmed by among others Bunge (1963, pp. 96-8), and Newton-Smith (1978, p. 77, and 1981, pp. 230-1). On the present treatment, of course, the thesis that the simplicity of theories is not indicative of those theories' observational success is subsumed under the thesis that simplicity is an aesthetic rather than logico-empirical property of theories.

uniform circular motions within explanations of observed planetary positions, and expressed displeasure at the Ptolemaic failure to adhere to this methodological tenet. J. Dalton wished chemical theory to be simple in supposing that the molecule of a chemical compound was formed of the smallest possible number of atoms of each of its constituent elements, e.g. that a molecule of water was composed of one atom of each of hydrogen and oxygen and a molecule of ammonia of one atom of each of hydrogen and nitrogen. E. Mach wished physical theory to be simple in appealing to only those physical entities of which the manipulation was necessary to account for observational data, and thus refusing to countenance the appeal to Newtonian absolute space and time. Lastly, Einstein wished physical theory to be simple in resting upon the smallest possible number of independent theoretical postulates. Implicitly described in the preceding paragraph are four forms of simplicity which a theory could be said to possess. The degree to which a theory possesses one of these forms is uncorrelated with the degree to which it possesses another. For instance, a physical theory which is highly simple in the sense desired by Einstein will not because of this be highly simple in the sense prescribed by Mach: a theory of which the number of independent deep postulates has been reduced to a minimum may well refer to a host of observationally-vacuous

---

7 Copernicus wrote in the Commentariolus (Swedlow ed. 1973, pp. 434-6; emphasis in the original): 'The theories concerning these matters that have been put forth far and wide by Ptolemy and most others [...] envisioned certain equant circles, on account of which it appeared that the planet never moves with uniform velocity either in its deferent sphere or with respect to its proper center. Therefore a theory of this kind seemed neither perfect enough nor sufficiently in accordance with reason. [...] I often pondered whether perhaps a more reasonable model composed of circles could be found from which every apparent irregularity would follow while everything itself moved uniformly, just as the principle of perfect motion requires.' A fuller discussion of the simplicity of the Copernican theory will follow in chapter 10.

8 Dalton's simplicity-criterion is examined in Bernatowicz (1970).

9 One of the most celebrated formulations of the Machian principle of simplicity is contained in his (1883), p. 586: 'Science [...] may be regarded as a minimal problem, consisting of the complete possible presentment of facts with the least possible expenditure of thought' (emphasis in the original).

10 Einstein left several formulations of his simplicity-criterion. He wrote for instance (1936, p. 352; emphasis in the original): 'The aim of science is, on the one hand, a comprehension, as complete as possible, of the connection between the sense experiences in their totality, and, on the other hand, the accomplishment of this aim by the use of a minimum of primary concepts and relations.' When commenting upon a discrepancy of up to ten per cent between the measured value of a gravitational deviation of a light-ray and the magnitude of the effect calculated from general relativity, he weighed structural simplicity against any empirical deficiency of the theory: 'For the expert, this thing is not particularly important, because the main significance of the theory does not lie in the verification of little effects, but rather in the great simplification of the theoretical basis of physics as a whole' (cited in Holton 1973, p. 236). Further discussion of Einstein's appeal to simplicity-criteria is offered in Hesse (1974), pp. 239-55, or Elkana (1982).
entities. Thus, mere mention of the degree to which a theory exhibits one or other of these forms of simplicity, accompanied by no specification of the form of simplicity to which this degree relates, will fail to constitute a complete characterization of the simplicity of that theory. What is required for a complete such characterization to be compiled is a specification of the degree to which the theory exhibits each of the possible forms of simplicity which may be envisaged in a theory. In other terms, an exhaustive characterization of the simplicity of a theory would tabulate the degree to which the theory exhibits ontological parsimony, the degree to which it exhibits moderation in the appeal to independent theoretical postulates, and so on.

Of course a stipulation of the simplicity which one would wish to see exhibited by theories in a certain domain of science need not be as encyclopaedic as a complete characterization of the simplicity actually exhibited by a theory. For whilst the latter codification must specify the degree to which the theory exhibits each of the forms of simplicity which may be envisaged in scientific theories, the former codification need stipulate only the degree to which theories ought to exhibit each of a limited number of forms of simplicity, viz., those forms to which one has resolved to accord particular value in theory-evaluation. Typically a normative codification of the simplicity of theories will specify – as did e.g. the codification of Mach or of Einstein described above – a single form of simplicity which one wishes preferentially to see embodied in theories: such a codification will hence be far less weighty than a complete specification of the simplicity of any actual theory.

Even if it is reduced in this way to its least terms, an exhaustive stipulation of the simplicity which one wishes to see embodied in theories must consist of two independent items of information: it must stipulate the form of the simplicity that theories ought to embody, and the degree to which they are to embody simplicity of that form. In consequence, two independent acts of judgement will be required in order to gauge the extent to which the simplicity of any given theory accords with a canon prescribing the simplicity which theories ought to possess: the first act of judgement will be needed to evaluate the extent of the agreement between the degree of the theory’s simplicity on the one hand and the degree of simplicity prescribed of theories by the canon on the other, the second to appraise the form of the simplicity exhibited by the theory by reference to the
form which the canon indicates that theories ought to possess.¹¹

In other words, an exhaustive assessment of the simplicity of a theory will require two independent evaluations to be made, and only one of these will be an evaluative measurement of the degree of the theory's simplicity: thus, one of the two evaluations which form part of a complete assessment of a theory's simplicity will consist of the application of a criterion other than the simplicity-criterion indicated by philosophers of science as being logico-empirical in nature, or correlated to the theory's likely future observational success.

The further suggestion of this section is that, while the evaluation of the degree of simplicity of a theory is an evaluation of that theory on logico-empirical grounds and hence a guide to its likely truth or empirical adequacy, the evaluation of its form of simplicity is an aesthetic judgement, or a judgement of which the verdict is not correlated with the theory's probable degree of future observational success. The claim that an appraisal of the form of simplicity possessed by a theory is independent of judgements of the likelihood that this theory should demonstrate empirical adequacy is supported by the following observation: that while the logic of inference suggests that a simpler theory is *ceteris paribus* more likely to demonstrate empirical adequacy than a less simple theory, it gives no indication of the form or forms of simplicity apt to maximize this likelihood. The logic of inference has no grounds upon which to tell whether a theory which exhibits e.g. ontological parsimony is *ceteris paribus* more or less likely to demonstrate empirical adequacy than a theory which exhibits moderation in the appeal to theoretical postulates. From among a set of theories exhibiting the form of simplicity selected by the evaluative canon which is applied to its domain of science, it is justified on logico-empirical grounds to choose for adoption the one which exhibits that form of simplicity to the highest degree; but the prior choice of which form or forms of simplicity one should like to see embodied in theories is taken on grounds other than logico-empirical.

This observation is evidence for the claim that evaluations of the form of

¹¹ There is ample historical evidence for the claim that scientists find themselves forced to express preferences between alternative forms of simplicity which theories may exhibit before the precept of choosing for adoption the simplest theory among several available ones can be carried out. For instance, as Holton (1978, p. 299, note 8; emphasis in the original) recounts, 'Einstein and Planck debated strongly in 1914 whether the simplest physics is one that regards as basic accelerated motion (as Einstein had come to believe) or unaccelerated motion (as Planck insisted).’ Clearly while this difference of opinion persisted, even if Einstein and Planck had concurred in following the methodological principle of each choosing to adopt that theory which exhibited to the greatest degree the form of simplicity which they wished to see embodied in theories, they would still have failed to agree on which theory to adopt.
simplicity exhibited by a theory are independent of empirical evaluations, and hence are aesthetic evaluations in the sense required in the present treatment.

4. Symmetry

Scientists have on many occasions evaluated theories for their possession of forms of symmetry. That the canonical formulation of Maxwell’s equations appears to reveal hidden symmetries in classical electrodynamics is frequently mentioned by physicists as a matter for aesthetic delight which serves to strengthen the attachment which they feel to the theory. H.G. Cassidy speaks of a pleasure of *prima facie* aesthetic nature which he obtained in perceiving a form of symmetry in the structure of a theory of which he conceived:

> While I was listening to a piano concerto, the idea suddenly occurred to me that it should be possible to prepare electron exchange polymers. I was at once certain that this was feasible, and I felt the fitness of the idea in complementing the already well-known proton exchange polymers. [...] Once the symmetry of the relationship became apparent to me I experienced great pleasure and excitement.  

Cassidy’s account of this episode suggests unambiguously that the perception of the form of symmetry in the theory which resulted from his idea served not only as stimulus to pursue the idea further in research, but also as initial grounds for the entrenchment of the theory in his system of beliefs, and hence as grounds upon which an early favourable evaluation of the theory was passed.

There exist many other historical examples of the appeal to forms of symmetry perceived in a theory as grounds for the adoption or preferment of that theory. At the heart of the arguments by which Einstein was led to doubt the validity of elements of classical physics and to extend initial support to the theory of special relativity lay considerations of the forms of symmetry which it was

---

12 Some of the aesthetic aspects of symmetry-considerations in theory-appraisal are discussed in Rosen (1975), pp. 120-2, and Zee (1986), *passim*.

13 Cassidy (1962), p. 57. Cassidy has recalled his appeal to symmetry-considerations in this episode also in an interview with Alexenberg contained in the latter’s (1981), pp. 147-8. In Cassidy (1966), pp. 185-6, the author suggests that his inference was rather an application of analogical reasoning.
appropriate to demand of physical theories.\textsuperscript{14}

Again, some time after Planck's theory of black-body radiation of 1900, the idea had gained currency that light exhibited corpuscular behaviour. In 1923 L. de Broglie suggested - 'purely on grounds of intellectual beauty', in the words of M. Polanyi\textsuperscript{15} - that material particles correspondingly possessed wave-like properties, advancing a relation between the momentum of a particle and the wavelength of the associated undulation which exactly mirrored Planck's equation linking the energy of the light-wave to its frequency. De Broglie's theory failed to gather empirical confirmation until 1927. In the years in which the physics community was unable to muster evidence of a logico-empirical kind sufficient to ensure the acceptance of the theory, the community's favourable attitude to it was motivated mainly by the observation that the theoretical corpus of elementary physics demonstrated a striking form of conceptual symmetry if the suggestion of de Broglie was incorporated into it.

The historiographic evidence presented thus far in this section gives some idea of the variety of the forms of symmetry of which the possession by theories has at one time or another of the history of science been deemed to be desirable. The existence of several forms of symmetry each of which it may be appropriate to prescribe of theories implies that an evaluative canon which makes mention of the feature-class of symmetry will typically indicate a preference for one particular form of symmetry: on the application of this canon in cases of theory-assessment, the theories which exhibit this form of symmetry will receive preference over others.

The decision to accord preference to theories which exhibit a particular form of symmetry over theories which exhibit other forms resembles the decision to prefer theories which \textit{ceteris paribus} demonstrate some particular form of simplicity: neither of these decisions - it is here suggested - is justifiable on logico-empirical grounds. There is no reason to believe that a sequence of theory-choices performed in accordance with a prescription \textit{ceteris paribus} to prefer theories exhibiting some particular form of symmetry would lead systematically to empirical success greater than that which would be attained through a series of choices which accorded preference to theories exhibiting any other form of symmetry.

\textsuperscript{14} A fuller discussion of the role of considerations of symmetry in Einstein's evaluations of classical physics and the theory of special relativity is contained in chapter 10.

\textsuperscript{15} Polanyi (1958), p. 148.
CHAPTER SIX: AESTHETIC FEATURES TO HAVE WON COMMENDATION

Far from constituting a reliable guide to choosing theories which demonstrate the greatest logico-empirical virtues, instructions to prefer some particular form of symmetry in acts of theory-choice appear mostly to confer a retrospective validation or blessing on a past happy decision – taken typically on grounds independent of considerations of symmetry – to adopt a theory which happened to exhibit that form of symmetry and which proceeded to demonstrate great empirical success. A scientific community’s canon of theory-evaluation appears to take note of the empirical fortune of such a decision by attributing in succeeding cases of theory-evaluation an increased weight to the form of symmetry which the theory chosen by that decision happened to exhibit.

Evidence that norms of preference for a particular form of symmetry fail to constitute a reliable guide to the logico-empirically most successful theories, and on the contrary amount merely to retrospective weighting of the symmetry-features of theories which have up to now demonstrated empirical success, derives from those historical cases in which norms of preference for a particular form of symmetry were revised in response to either an empirical failure of a theory which exhibited that form of symmetry or the rise of an empirically-successful theory which possessed some new form of symmetry. M.L.G. Redhead notes one of the occasions on which the precept that preference in theory-choice be given to theories demonstrating a certain form of symmetry was altered in response to a decrease in the apparent ability of some theories exhibiting that form of symmetry to account for experimental data:

The discovery that in the weak interactions of elementary particles nature provides an ultimate distinction between right-handed and left-handed reference frames reminds us that the superlaws of symmetry, as Wigner calls them, are as liable to empirical revision as other laws of physics having a less obviously intuitive character.\(^\text{16}\)

Even the norms prescribing that preference be given to theories which exhibit the most intuitively appealing forms of symmetry are promptly disavowed if theories which embody those forms cease scoring empirical successes, or if a theory which violates those forms demonstrates great empirical virtues. The fact that a community’s preference for certain forms of symmetry appears to await demonstrations of empirical success by theories which exhibit those forms, rather than themselves reliably leading to empirically-successful theories, helps to locate symmetry-criteria among the aesthetic rather than the logico-empirical criteria of

\(^{16}\) Redhead (1975), p. 105.
CHAPTER SIX: AESTHETIC FEATURES TO HAVE WON COMMENDATION

theory-assessment of scientific communities.17

5. Analogical tractability

The observation that it is possible to draw relations of analogy between the structure or claims of a theory under evaluation and those of a particular other (typically, better-confirmed or better-entrenched) theory enters among the grounds upon which a community may be drawn to lend support to the new theory. On many occasions a theory has either gathered endorsement in virtue of the fact that it supported relations of analogy with a particular other theory, or repelled members of its community by proving unsusceptible to tractability in terms of some analogy which the community has been led to value. The question arises, what is the nature of the evaluative grounds to which one appeals when one argues that a theory ought to be endorsed in virtue of the fact that it exhibits the capacity to support a relation of analogy with a particular other body of theory?

The 1960s produced some treatments of analogy in science which were less dismissive of this device than positivist philosophy had been.18 These more modern treatments were in the main concerned to establish that models constructed on the basis of analogies between two or more theories in different domains of science cannot be reduced to only the role of suggestion or stimulus to theory-formulation attributed to them by positivist philosophy of science, and carry weight in theory-evaluation also. Such new treatments concentrated almost entirely on the logico-empirical aspects of scientific models, however, and withheld attention from the feeling of aesthetic satisfaction which commonly derives from the perception of analogical relations between, say, theories in different domains of science.

That the discovery of the existence of a possible analogy is an experience which carries aesthetic aspects can hardly be in doubt. Analogical reasoning aims to reveal the presence in different constructs of common elements, a typically aesthetic concern. As a consequence, the devices of metaphor and analogy are

17 The susceptibility of the symmetry-criteria of a community to revision in the light of the degree of empirical success won by theories demonstrating various forms of symmetry will be studied and explained in the next chapter, which attributes a metainductive origin to aesthetic canons of theory-assessment.

possible sources of aesthetic pleasure in science no less than in literary or artistic
endeavours.¹⁹

The present treatment suggests that to evaluate a theory in virtue of the
fact that its structure or claims can be analogically related to those of some
particular other theory constitutes an aesthetic evaluation of that theory. Of
course, within the present treatment to advance this claim is chiefly to suggest
that the degree to which a theory supports analogical relations with some
particular other theory is not correlated with the degree to which it possesses
logico-empirical virtues. The truth of this observation may be suggested by a
study of the development of quantum mechanics from 1913 to 1927, a period
which witnessed the loss and recovery of analogical tractability on the part of the
leading theories in the field.²⁰

In 1913 Bohr proposed a model of the atom which retained its conventional
visualization as a miniature planetary system. The subsequent decade revealed
however the incapacity of this model to describe atoms more complex than the
simplest, that of hydrogen. In 1925 Heisenberg originated a new version of
quantum mechanics, matrix mechanics, couched in a mathematical formalism
uninterpreted by models and referring throughout to particles of an
unvisualizable nature: in this succession the concept of electron had evolved from
analogue of a miniature billiard ball to purely abstract entity. Heisenberg asserted
that he found this formal approach congenial to his nonvisual mode of thought;
Schrödinger was on the contrary disappointed by the lack of visualizability in the
theory and in 1926-27, partly in reaction to Heisenberg’s work, developed the
theory of wave mechanics. This theory was logically equivalent to Heisenberg’s
but pictured subatomic particles as wave-packets. Schrödinger described the
genesis of his theory and his reaction to Heisenberg’s results:

My theory was inspired by L. de Broglie [...] and by short but
incomplete remarks by A. Einstein [...]. No genetic relation whatever
with Heisenberg is known to me. I knew of his theory, of course,
but felt discouraged not to say repelled, by the methods of
transcendental algebra, which appeared very difficult to me and by

¹⁹ A study of the features common to the use of analogy in science and in literature is
Beer (1983), pp. 79-103. Investigation of the similarities between the use of analogy
in art and in science has prompted attention for the separate finding that analogies
of a literary type feature in some scientific reasoning; the literary metaphors of
Darwin, for instance, are studied by Hyman (1959), pp. 14-78, and Gruber (1978).

²⁰ For fuller details of these episodes in the development of quantum mechanics see
Miller (1984), pp. 125-83, on which the following account draws.
the lack of visualizability.\textsuperscript{21}

What did Schrödinger and many other physicists between his time and the present find displeasing in Heisenberg’s matrix mechanics and appealing in Schrödinger’s own wave mechanics? The reason for the differential evaluation cannot have been the greater observational success of the latter theory, or any logico-empirical feature of it, for the two theories are logically equivalent and consequently equally well supported by any evidence. The reason for the preference must thus have been independent of considerations of observational success or of logico-empirical features.

In short, the preference of Schrödinger for his own theory over Heisenberg’s must have been motivated by non-utilitarian considerations. It must have been the expression of an aesthetic canon of theory-assessment, one which was not tied to appraisals of the degree of possible utility of the competing theories. It is easy to see what feature of wave mechanics, lacking in matrix mechanics, is capable of triggering this non-utilitarian preference: its analogical visualizability. On this argument, the requirement that a theory should demonstrate analogical tractability in terms of a given model is a \textit{desideratum} introduced by an aesthetic canon of appraisal of scientific theories.

In discussing the requirement that scientific theories should demonstrate analogical tractability, attention should be paid to the distinction drawn in section 1 of this chapter between aesthetic features and aesthetic feature-classes. It was pointed out there that on this distinction only aesthetic features, and not feature-classes, were an appropriate object of stipulation by aesthetic canons of theory-appraisal; in the application of this distinction to the discussion of simplicity, for instance, it was remarked that a determinate aesthetic evaluative canon could legitimately stipulate preference for theories that exhibit a certain form of simplicity but not for theories that are generically ‘simple’. The same caveat applies to the discussion of analogical tractability. An aesthetic evaluative canon which demands that theories demonstrate a generic quality of ‘analogical tractability’ would not possess sufficient determinacy to be applied in cases of theory-choice. Since one may erect an analogy between some aspect or other of

\textsuperscript{21} Quoted \textit{ibid.}, p. 143. Heisenberg’s own account of the rise of wave mechanics is contained in his (1971), pp. 70-3. He appears to agree that wave mechanics possessed some aesthetic appeal not matched by other formulations of quantum mechanics: ‘Schrödinger first of all explained the mathematical principles of wave mechanics by using the hydrogen atom as an illustration. All of us were delighted to see his elegant and simple solution by conventional methods of a problem that Wolfgang Pauli had been able to solve only with great difficulty using quantum mechanics’ (\textit{ibid.}, pp. 72-3). The meeting here described was held in 1926.
any or virtually any theory under evaluation on the one hand and some suitably-chosen other theory or model on the other, the application of such a generic canon would not exercise a determinate influence on theory-choice. An aesthetic evaluative canon applied in scientific practice may therefore contain the demand that theories should support analogical relations with a set of theories of a certain specified kind, but it would be pointless for such a canon to issue a request that theories should exhibit generic ‘analogical tractability’. In terms of the distinction referred to above, this is because ‘analogical tractability’ is an aesthetic feature-class: just as ‘such-and-such a form of simplicity’ is an aesthetic feature which belongs to the feature-class of simplicity, ‘tractability in terms of such-and-such an analogy’ is a feature which belongs to the feature-class of analogical tractability.

Thus, the aesthetic evaluative canon of a community will stipulate that theories in a certain domain of science exhibit tractability in terms of analogy with a specified model or family of models. When Lord Kelvin insisted, ‘I never satisfy myself until I can make a mechanical model of a thing’;22 intending to apply this test in his evaluation of the acceptability of newly-formulated scientific theories, he expressed one of the requirements contained in the canon for the aesthetic evaluation of physical theories which ruled at his time. The requirement was for theories in physics to exhibit tractability in terms of mechanistic models, or by analogy with the previous and well-entrenched theories of Newtonian mechanics.

Clearly the imposition of this requirement upon physical theories will not be thought appropriate at all times in the history of science: it would be deemed inappropriate, for instance, to require mechanistic interpretability of theories in quantum physics today. Thus it is seen that the set of models or of analogical relations which it is thought appropriate to require of theories changes in time. The process by which this succession occurs will form the subject of the next chapter, but some initial observations about the manner in which the favour of the scientific community passes from one model or family of models to another may be reported here.

Theory-choice in anatomy since at least the sixteenth century has been influenced by an evaluative canon which has prescribed that preference be attributed to anatomical theories which exhibited relations of analogy with theories in other domains of science, and frequently with theories in the better-entrenched physical sciences. But the set of theories which were considered the

---

most appropriate sources for these analogies has changed in time; these changes can be interpreted as revisions of the prescriptive content of the evaluative canons applied to theory-choice in anatomy. Hesse attributes to N. Wiener the observation that

there have been three stages in the scientific description of human beings according to what was the most typical machine in use during the period - first, in the seventeenth and eighteenth centuries, clockwork mechanisms described by analogies from dynamics; then in the nineteenth century, heat engines described by analogies from thermodynamics; and now communication devices described by analogies from electronics.23

As the next chapter will argue in greater detail, this succession may be explained on the assumption that the scientific community at each moment chose as the most suitable source of models in anatomy theories of the domain in the physical sciences which had in the immediately preceding period demonstrated the most eye-catching empirical success. The observationally most successful theories available in the seventeenth and eighteenth centuries were Cartesian and later Newtonian mechanics, both of which looked to analogies with clockwork mechanisms in their description of dynamic phenomena such as planetary systems: such mechanistic models gathered over time great weight in the community's stock of models, and it is of no wonder that this period looked to interpret newly-formulated theories in their terms and to evaluate the worth of new theories according to whether they admitted or failed to admit such interpretations. By the nineteenth century the glory of the observational success of Newtonian mechanics had to some extent been eclipsed by that of theories in thermodynamics: the community's stock of favoured models was updated to take account of the new weighting. In the twentieth century the empirical success of solid-state physics has similarly redefined the stock. The aesthetic evaluative canon of the community evolved in keeping with these changes.24

23 Hesse (1954), p. 140; Hesse refers the reader to Wiener (1948), pp. 39-40, as well as to Young (1951), pp. 24-38, where a similar suggestion is advanced. Hanson mentions an analogous succession of models (1971, pp. 77-8): 'the Saturn-model of the hydrogen atom, the shell-model of the atomic nucleus, the telephone switchboard-model of the human brain, the hydraulic conduit-model of the neural fibers.'

24 This account may be read as an explication of Hesse's remark that 'contemporary categories of thought and social conditions may take a hand in moulding [...] scientific analogies' (1954, p. 140).
CHAPTER SIX: AESTHETIC FEATURES TO HAVE WON COMMENDATION

6. Consonance with metaphysical presuppositions

A scientific theory of more than technical interest will be said to exhibit consistency with the members of a certain set of metaphysical presuppositions and to conflict with the members of another set of similar presuppositions. It is in the nature of a scientific community to express endorsement, even if only tacitly, to a body of metaphysical presupposition. Thus among the grounds upon which the community evaluates and chooses between competing scientific theories is the degree of consistency between the theories proposed and the set of presuppositions to which the community pledges allegiance. For instance, the part of the natural-philosophical community which in the late seventeenth and early eighteenth centuries professed allegiance to Cartesianism opposed Newton’s theory of gravitation in virtue of its appeal to attraction at a distance, which appeared inconsistent with the Cartesian metaphysical tenet that space consisted of a plenum of corpuscles interacting solely by collision.25

What is the nature of the grounds to which the community appeals when it extends or denies support to a scientific theory in virtue of the degree of its consistency with some metaphysical presupposition? It has not hitherto been customary for philosophers of science to describe these grounds as aesthetic in nature; indeed, many philosophers who have considered the appeal in scientific methodology to aesthetic arguments have tended on the contrary to number aesthetic grounds of theory-evaluation within a more general category of metaphysical grounds.26 Yet the property of exhibiting consistency with metaphysical presuppositions resembles the other features of theories described in this chapter to which appeal is made in theory-evaluation. First, the perception of a consonance or harmony between one’s metaphysical presuppositions on the one hand and a scientific theory with which one has been acquainted on the other affords a gratification of prima facie aesthetic kind. Secondly, since – presumably as a matter of definition – there is no empirical warrant for one’s metaphysical presuppositions, the degree of a theory’s consistency with that set of presuppositions is uncorrelated with the degree to which that theory possesses empirical virtues. For instance, the consistency of the Newtonian theory of gravitation with the denial of validity in metaphysics to Cartesian corpuscularism

25 The Cartesians’ and Leibniz’s attack on action at a distance is discussed in Hesse (1961), pp. 157-63.

26 Some of the works in philosophy of science which subsume aesthetic grounds for theory-evaluation within a more general category of metaphysical grounds were mentioned in chapter 2 above.
does not make Newton’s theory more likely to demonstrate observational success than a theory which, like Cartesian cosmological theory, is consistent with Cartesian metaphysical doctrine: what in point of fact allowed Newton’s theory to attain a measure of observational success greater than that won by Descartes’s theory was the difference between the properties attributed to natural agencies by the detailed mathematical claims of the two theories. Thirdly – and this observation will be investigated in greater detail in chapters to follow – the set of metaphysical presuppositions to which scientific communities have lent their assent appears to have varied in history in a manner very similar to that of the variation of the set of analogical relations, forms of symmetry, or forms of simplicity to which communities have attributed weight. That is to say, scientific communities appear at each moment to have defined the set of metaphysical presuppositions to which to lend assent in part so as to include those presuppositions with which the empirically most successful theories of the immediately preceding period exhibited consistency. It will be suggested in the next chapter that to exhibit this form of historical mutability is one of the characteristics of aesthetic criteria of theory-evaluation.27

27 Chapter 10 will contain an extended study of an instance of the evaluation of a theory in virtue of the degree of its consistency with a metaphysical presupposition: it will outline Einstein’s reaction to quantum mechanics, which was hostile in consequence of Einstein’s allegiance to – and the theory’s violation of – the metaphysical requirement that theories should demonstrate causal determinism. There will also be presented further evidence for the claim that the requirement that theories accord with metaphysical presuppositions is aesthetic in nature.
Chapter Seven
THE INDUCTIVE CONSTRUCTION
OF AESTHETIC PREFERENCE

1. The determinacy of theory-choice

Evaluations of a given theory on empirical criteria and on aesthetic criteria are according to the analysis of chapter 5 in principle independent of one another: in a situation demanding an act of theory-choice the preference indicated by the application of empirical criteria will in general differ from that reached by the application of aesthetic criteria. A theory's score on, so to speak, the axis of empirical adequacy has in principle no systematic correlation with its score on the axis of perceived aesthetic virtue.

The independence of the two canons of evaluative criteria offers both benefits and dangers to the scientist. The benefit which accrues from possessing a double canon of appraisal of scientific theories is the opportunity that a narrower and more discriminating ideal may thereby be constructed of the object of research, if – as occurs in this case – the two evaluative canons stipulate complementary qualities of theories to be desirable. This opportunity is envisaged by those methodologists who would by this means avert the underdetermination of theory-choice. At the same time however the very fact that the two canons of evaluative criteria prescribe different qualities to be desirable in scientific theories opens the possibility that instances of theory-choice should witness a conflict between the recommendations yielded by the two classes of criteria, and thus remain undecided.

The prospect of a similar methodological dilemma is not peculiar to models of scientific rationality which like the present one envisage two sets of criteria of theory-evaluation purporting to refer to two different kinds of features of scientific theories: the danger of conflict between evaluative criteria haunts any canon which comprises more than one independent criterion. The same dilemma was for instance recognized by P. Frank:
CHAPTER SEVEN: THE INDUCTIVE CONSTRUCTION OF AESTHETIC PREFERENCE

It has been made clear by now that the requirements for the acceptance of a theory by scientists in the modern sense are 'agreement with observation' and 'simplicity.' [...] Which of them is the more important one? [...] If we have to choose between a theory that is in agreement with the facts but is very complicated, and a theory that is much simpler but does not agree as well in all details with the facts, which theory are we to choose?¹

The concern caused by the possibility of a conflict in theory-assessment between two criteria which are both members of a unified logico-empirical canon of evaluative criteria – such as Frank's 'agreement with observation' and 'simplicity' – is however lessened by the realization that those who formulated such a canon intended presumably that its repeated application in theory-choice should converge, or lead towards a single well-defined end-point of science, such as the realist's the complete true explanatory account of the universe. The uniqueness of the intended end-point of the process of repeated theory-choice encourages one to suppose that any divergence of the recommendations of two criteria within the canon will be recomposed when the canon is applied at a later instance of theory-choice: at that later stage, one may suppose, the theory which will impose itself to the choice of the community will combine within itself the separate virtues which caused the recommendations of the criteria of theory-choice to diverge at earlier stages.

Concern at the possibility of a conflict between different criteria of theory-choice is on the contrary more severe when it occurs between two sets of criteria which are not defined by reference to the same end-point of theory-choice, for there is then no prima facie expectation that the agreement between the criteria will increase as science nears its prescribed end-point. Such is of course the uncertainty encountered by the present model of theory-evaluation: while one of the classes of criteria which it envisages – the logico-empirical canon – aims at the end-point of empirical adequacy much as does Frank's canon of theory-

¹ Frank (1957), pp. 352-3. Frank answers his own question in two attempts. First he suggests that 'if we ask a scientist, he will probably answer that the decisive point is the agreement with the observed facts and that "simplicity" is of secondary importance' (ibid., p. 353). Upon further consideration he settles on the fence: 'the acceptance of a theory is always the result of a compromise between the requirement of "agreement with facts" and of "simplicity"' (ibid.). Of course the problem – which Frank fails to address, and which will be treated below – is to determine how the compromise between the requirements of different evaluative criteria is, or ought to be, struck by scientific practice. Frank's two-criterion model of theory-appraisal is, incidentally, enriched a little later in his treatment: 'there are actually three requirements that have been admitted by scientists: agreement with observations, simplicity, and agreement with common-sense experience' (ibid.). Frank's model of scientific criteria of theory-assessment nonetheless remains hopelessly simplistic.
CHAPTER SEVEN: THE INDUCTIVE CONSTRUCTION OF AESTHETIC PREFERENCE

assessment, the other class — that of aesthetic criteria — does not share that endpoint. This is because the aesthetic criteria express preferences among theories not in regard to their logico-empirical virtues but in regard to virtues which are defined as non-utilitarian and therefore non-empirical. Hence this account of theory-evaluation cannot hope to rule out the prospect of serious conflict between evaluative criteria by the route which is on the contrary open to models like that of Frank.

While the independence of the two sets of criteria raises the possibility that instances of theory-choice may remain undecidable, this danger does not necessarily arise. There are two cases in which this eventuality does not in fact occur. The first obtains if, notwithstanding the independence in principle of the two sets of criteria, some correlation should subsist between them in practice. The second would occur if a scientific community had disposed a procedure for the arbitration of any conflict between the two sets of criteria of evaluation, perhaps by stipulating the priority of one set over the other in any case of dispute.

Much of the remainder of this treatment will aim to show that scientific communities use both these methodological devices at different phases in the history of science to ensure that they retain the capacity to make univocal choices among theories: in some historical phases scientific communities maintain a correlation between the preferences expressed by their two classes of evaluative criteria, while at other times they confer privilege to the recommendations yielded by one of the two classes. The conditions under which a scientific community turns to one or the other of these ways of resolving the issue will define the nature of the two phases of scientific history.

2. The origin and mutability of methodological precepts

The model of the methodological history of science to be enunciated exploits the familiar distinction between the doctrines of metarationalism and metainductivism. These doctrines express views on the nature and development of methodological precepts in science. According to metarationalism, the norms of scientific methodology are formulated \textit{a priori} by inferences from the nature and goals of science: on this foundationalist account a methodological precept correctly inferred will remain forever valid. For instance, Popper believes that the
norms of refutationalism are inferred from the nature of science, and in particular from considerations on the impossibility of inductive validation and the asymmetry between confirmation and refutation; if this inference is valid, the norms of refutationalism possess necessity and irreplaceability in the scientific methodological canon. Metainductivism holds on the contrary that the norms of scientific methodology are developed and refined by an induction over those amongst all past proposed norms which have demonstrated the most fruitful applicability, so that the battery of norms to which a community adheres may evolve in response to a perceived inferiority of its performance relative to that of an alternative proposed battery.

One may initially harbour the hope of adjudicating between these two doctrines on historical evidence by gauging the degree of variability which has been exhibited over time by the methodological norms actually prescribed and followed in science. Metarationalism would naturally anticipate no variability in the substance of those norms, whilst perhaps accepting that a scientific community may inductively learn to formulate them or come progressively to acknowledge them; metainductivism trusts on the contrary to perceive in the history of science a succession of inductive changes in the norms endorsed by scientific communities.

The decisiveness of this historiographic test is unfortunately blunted by the fact that there exists wide disagreement in the interpretation of the historical record on this issue: different observers appear to perceive different degrees of historical variability in the methodological canons of science. At one extreme I. Scheffler sees in the historical record little or no methodological variability:

Underlying historical changes of theory, there is [...] a constancy of logic and method, which unifies each scientific age with that which preceded it and with that which is yet to follow.

Such constancy comprises not merely the canons of formal deduction, but also those criteria by which hypotheses are confronted with the test of experience and subjected to comparative evaluation.2

Clearly this reading of the historical record would lend no assistance to the metainductivist supposition that the history of science exhibits significant changes

---

2 Scheffler (1967), pp. 9-10. Clark (1962, pp. 103-4; emphasis in the original) expresses a similar conviction: 'Science has already achieved such a sufficiently clear self-conscious awareness of the intellectual structure of its own enterprise that while future operations will surely disclose new and unexpected results through novel and unpredictable techniques, the logical pattern of their production must and will remain basically the same as those of today.'
in procedural canons.\(^3\) A quite different historiographic finding is reported by L. Laudan, who criticizes the approach of those who like Scheffler view the history of science as methodologically changeless:

In some cases, proponents of such an approach have gone so far as to claim that all the actual standards of rational appraisal have remained constant through time. [...] We need waste little time on this approach. Virtually all the scholarly literature on the history of methodology shows unambiguously that such components of rational appraisal as criteria of explanation, views about scientific testing, beliefs about the methods of inductive inference and the like have undergone enormous transformations.\(^4\)

The perception of such transformations in science’s procedural canons casts doubt on the metarationalist account of the origin of such canons and offers evidential support to the metainductivist conception, in as much as this doctrine is one of the possible range of accounts able to predict and explain variability of methodology.\(^5\)

If the historical data concerning the degree of variability of methodological norms are susceptible to interpretations as different as those of Scheffler and Laudan, it becomes difficult to adjudicate between metarationalism and metainductivism as accounts of the genesis of norms of scientific procedure. In

---

\(^3\) This reading of the historical record has commanded much wider assent in philosophy of science than has the contrary view: indeed Laudan (1979, p. 45) states that ‘virtually every major philosopher of science, from Aristotle […] to Carnap […] has assumed that there is a set of rules for scientific method which are permanent and trans-temporal.’ Laudan cites as representative of this stance Carnap’s (1928).

\(^4\) Laudan (1977), p. 129; emphasis in the original. Laudan’s criticism is prompted by the same passage of Scheffler (1967) which is quoted above. J.R. Brown (1985, p. 298) suggests that this thoroughgoing attribution of variability to methodological norms sits uneasily with the main concerns of Laudan’s (1977), which appears to embrace the view that at least some components of science’s normative methodology are eternal or extrahistorical. This interpretative point does not detract from Laudan’s passage cited here as a statement of the position that methodological norms have varied in time, and to this extent as an illustration of the reading of the historical record which supports the metainductivist account of the genesis of methodology.

\(^5\) Laudan’s view that scientific methodology has exhibited a pronounced variability in history finds an extreme echo in Feyerabend’s writing, which of course draws also the normative conclusion (not pursued by Laudan) that for this reason it is deleterious to science to prescribe any subset of the widest potential range of methodological norms to the exclusion of any other subset. ‘The idea that science can, and should, be run according to fixed and universal rules, is both unrealistic and pernicious. It is unrealistic, for it takes too simple a view of the talents of men and of the circumstances which encourage, or cause, their development. And it is pernicious for the attempt to enforce the rules is bound to increase our professional qualifications at the expense of our humanity. In addition, the idea is detrimental to science, for it neglects the complex physical and historical conditions which influence scientific change’ (1975, pp. 295-6; emphasis in the original). In the previous chapter of that work Feyerabend reports case studies which, he believes, militate against the attribution of universal validity to any methodological rule.
in procedural canons. A quite different historiographic finding is reported by L. Laudan, who criticizes the approach of those who like Scheffler view the history of science as methodologically changeless:

In some cases, proponents of such an approach have gone so far as to claim that all the actual standards of rational appraisal have remained constant through time. [...] We need waste little time on this approach. Virtually all the scholarly literature on the history of methodology shows unambiguously that such components of rational appraisal as criteria of explanation, views about scientific testing, beliefs about the methods of inductive inference and the like have undergone enormous transformations.

The perception of such transformations in science's procedural canons casts doubt on the metarationalist account of the origin of such canons and offers evidential support to the metainductivist conception, in as much as this doctrine is one of the possible range of accounts able to predict and explain variability of methodology.

If the historical data concerning the degree of variability of methodological norms are susceptible to interpretations as different as those of Scheffler and Laudan, it becomes difficult to adjudicate between metarationalism and metainductivism as accounts of the genesis of norms of scientific procedure. In

---

3 This reading of the historical record has commanded much wider assent in philosophy of science than has the contrary view: indeed Laudan (1979, p. 45) states that 'virtually every major philosopher of science, from Aristotle [...] to Carnap [...] has assumed that there is a set of rules for scientific method which are permanent and trans-temporal.' Laudan cites as representative of this stance Carnap's (1928).

4 Laudan (1977), p. 129; emphasis in the original. Laudan's criticism is prompted by the same passage of Scheffler (1967) which is quoted above. J.R. Brown (1985, p. 298) suggests that this thoroughgoing attribution of variability to methodological norms sits uneasily with the main concerns of Laudan's (1977), which appears to embrace the view that at least some components of science's normative methodology are eternal or extrahistorical. This interpretative point does not detract from Laudan's passage cited here as a statement of the position that methodological norms have varied in time, and to this extent as an illustration of the reading of the historical record which supports the metainductivist account of the genesis of methodology.

5 Laudan's view that scientific methodology has exhibited a pronounced variability in history finds an extreme echo in Feyerabend's writing, which of course draws also the normative conclusion (not pursued by Laudan) that for this reason it is deleterious to science to prescribe any subset of the widest potential range of methodological norms to the exclusion of any other subset: The idea that science can, and should, be run according to fixed and universal rules, is both unrealistic and pernicious. It is unrealistic, for it takes too simple a view of the talents of man and of the circumstances which encourage, or cause, their development. And it is pernicious for the attempt to enforce the rules is bound to increase our professional qualifications at the expense of our humanity. In addition, the idea is detrimental to science, for it neglects the complex physical and historical conditions which influence scientific change (1975, pp. 295-6; emphasis in the original). In the previous chapter of that work Feyerabend reports case studies which, he believes, militate against the attribution of universal validity to any methodological rule.
in procedural canons. A quite different historiographic finding is reported by L. Laudan, who criticizes the approach of those who like Scheffler view the history of science as methodologically changeless:

In some cases, proponents of such an approach have gone so far as to claim that all the actual standards of rational appraisal have remained constant through time. [...] We need waste little time on this approach. Virtually all the scholarly literature on the history of methodology shows unambiguously that such components of rational appraisal as criteria of explanation, views about scientific testing, beliefs about the methods of inductive inference and the like have undergone enormous transformations.

The perception of such transformations in science’s procedural canons casts doubt on the metarationalist account of the origin of such canons and offers evidential support to the metainductivist conception, in as much as this doctrine is one of the possible range of accounts able to predict and explain variability of methodology.

If the historical data concerning the degree of variability of methodological norms are susceptible to interpretations as different as those of Scheffler and Laudan, it becomes difficult to adjudicate between metarationalism and metainductivism as accounts of the genesis of norms of scientific procedure. In

---

3 This reading of the historical record has commanded much wider assent in philosophy of science than has the contrary view: indeed Laudan (1979, p. 45) states that ‘virtually every major philosopher of science, from Aristotle [...] to Carnap [...] has assumed that there is a set of rules for scientific method which are permanent and trans-temporal.’ Laudan cites as representative of this stance Carnap’s (1928).

4 Laudan (1977), p. 129; emphasis in the original. Laudan’s criticism is prompted by the same passage of Scheffler (1967) which is quoted above. J.R. Brown (1985, p. 298) suggests that this thoroughgoing attribution of variability to methodological norms sits uneasily with the main concerns of Laudan’s (1977), which appears to embrace the view that at least some components of science’s normative methodology are eternal or extrahistorical. This interpretative point does not detract from Laudan’s passage cited here as a statement of the position that methodological norms have varied in time, and to this extent as an illustration of the reading of the historical record which supports the metainductivist account of the genesis of methodology.

5 Laudan’s view that scientific methodology has exhibited a pronounced variability in history finds an extreme echo in Feyerabend’s writing, which of course draws also the normative conclusion (not pursued by Laudan) that for this reason it is deleterious to science to prescribe any subset of the widest potential range of methodological norms to the exclusion of any other subset: ‘The idea that science can, and should, be run according to fixed and universal rules, is both unrealistic and pernicious. It is unrealistic, for it takes too simple a view of the talents of man and of the circumstances which encourage, or cause, their development. And it is pernicious for the attempt to enforce the rules is bound to increase our professional qualifications at the expense of our humanity. In addition, the idea is detrimental to science, for it neglects the complex physical and historical conditions which influence scientific change’ (1975, pp. 295-6; emphasis in the original). In the previous chapter of that work Feyerabend reports case studies which, he believes, militate against the attribution of universal validity to any methodological rule.
these circumstances the suggestion dawns that perhaps each of these models is capable of capturing certain aspects of the process of construction of scientific communities’ methodological canons, so that the way in which communities formulate their methodological canons would best be modelled by a combination of the two accounts. This conclusion would explain how it is that history yields evidence which lends support to both the metarationalist and the metainductivist models of method-construction.

The question then arises of how these two models can suitably be combined in a coherent view of the origin of methodological norms. Some authors appear determined to pursue the synthesis by leaving their models of scientific rationality in ambiguity, suggesting in one context that a methodological canon is inferred metarationalistically from the perceived goal of scientific research, and in another context that the authority of each criterion is increased by the inductive observation of its fruitfulness in research. For instance Newton-Smith, surveying the ‘factors [...] which can serve as fallible indicators of likely long-term observational success’, asserts that all criteria of theory-choice are formulated or selected by an inductive procedure:

The grounds for including any particular factor will be meta-inductive. If we can locate factors that have guided scientists in making theory choices which turned out to be correct on the ultimate test, we shall have inductive grounds for operating within the constraints of these particular inductive factors.6

Other passages of the same treatment however appear devoted to undermining this view. For instance, having allowed to methodological precepts a variability occasioned by the community’s inductive reassessment of their effectiveness, Newton-Smith minimizes their rate of inductive change – and loses to his model of scientific rationality some historiographic flexibility – in averring that ‘there does seem to be considerable consistency in what the scientific community in different cultures and different ages holds to be the good-making qualities of a theory.’7 Newton-Smith further puzzles the reader by appending to the opinions just outlined the further statement that methodological precepts are formulated by scientific communities through inferences from the goal of science, an origin which appears incompatible with the community’s progressively refining its

7 Ibid., p. 112.
methodological criteria by inductively examining their effectiveness.\footnote{\textit{Ibid.}, pp. 223-4. Newton-Smith's account of the steps of the inference from a statement of the goal or nature of science to a methodological canon was examined in chapter 2 above.}

The ambivalence of judgements of the degree of variability of methodological precepts in history, and the consequent difficulty of adjudicating between metarationalist and metainductivist accounts of the origin of those precepts, merit a response possessed of more internal consistency than one which seemingly attributes the origin of one and the same set of methodological norms to two different mechanisms. If, as appears from the historiographic literature, both metarationalist and metainductivist accounts of the origin of methodological norms seem to capture part of the truth, there is a superior method by which to incorporate the insights of both these views into the same model of scientific rationality.

Epistemological theorists generally posit that the scientific community's complex of methodological norms has \textit{in toto} either demonstrated past fixedness and thus supported a metarationalist view, or undergone evolution and therefore been appropriately ascribed to a metainductivist origin. However there is no reason why both these modes of method-construction should not operate simultaneously in scientific history on separate categories of methodological norms. The present treatment, which has already and on independent grounds established two differently-defined classes of norms of theory-evaluation, is particularly well placed to pursue this suggestion. The thesis that shall be expounded here is that whereas empirical criteria of theory-assessment are formulated by inferences from first principles characteristic of metarationalism, aesthetic criteria are constructed \textit{a posteriori} by metainductivism.

\section*{3. The aesthetic metainduction}

Metarationalism is clearly responsible for the genesis of empirical criteria because, on the account of chapter 2, their inclusion among the \textit{desiderata} of theories derives entirely from the \textit{a priori} definition of the goal of science, the complete and empirically adequate explanatory account of the universe. The requirements of internal consistency or predictive accuracy are prized not because they have
previously, inductively been witnessed to accompany empirical adequacy but because they are the elements of an explication of what it is for a theory to be empirically adequate. Empirical criteria in other terms do not provide an ampliative connotation of empirical success, of which the correlation with empirical success might await discovery by inductive enumeration: on the contrary, they are elements deemed constitutive of empirical success, in the sense that for a theory to achieve empirical success is just for it to satisfy the requirements of the logico-empirical criteria of evaluation. Therefore the formulation of the logico-empirical criteria results from a judgement about the nature of empirical success, and is achieved metarationalistically or without inductive recourse to the history of science. It remains of course possible for formulations of empirical criteria of theory-assessment to be inductively refined by a scientific community but this is irrelevant to the a priori logical status of such criteria with respect to scientific practice.

While the attribution of a metarationalist origin to logico-empirical criteria of theory-assessment follows naturally if one understands them to prescribe to theories features constitutive of the notion of empirical success, aesthetic evaluative criteria possess no such link to the goal of science. Hence, the present treatment will not attribute a metarationalist origin to aesthetic evaluative criteria as well as to logico-empirical criteria. On the contrary, the remainder of this section will be devoted to outlining a metainductivist account of the origin within scientific communities of aesthetic canons of theory-assessment.

In order to render intuitively more acceptable such an attribution, consider first some of the characteristics of aesthetic evaluations of scientific theories. Such evaluations, as the previous chapter has illustrated, have demonstrated considerable variability in time. At some stage in the history of science certain aesthetic features may have been seen as desirable in theories, and a particular theory's possession of those features may have formed part of the case for its adoption by the community; but there is no guarantee that those aesthetic features will be seen as desirable in theories at other moments in the history of science, either before or after the episode mentioned. On the contrary, at some other time in history a theory's possession of the same aesthetic features may be seen as immaterial or damaging to the case for its adoption by the community. In brief, an aesthetic feature – call it feature i – which is present in some theories may be seen at different times in history as weighing in favour or in disfavour of the adoption of theories which exhibit i. Since the aesthetic feature specified as i upon which such variable judgements are passed is not of itself liable to change,
the sole entity which by its changes may be held responsible for the changing judgements is the community's canon of aesthetic evaluative criteria. The fact that a theory's possession of a certain aesthetic feature may be seen as first to the theory's credit and then to its discredit would on this assumption be explained by a change in the degree of favour attributed to that feature by the community's aesthetic evaluative canon.

The historiographic contention that communities' aesthetic evaluative canons change in time, to yield at different times different evaluations of the same aesthetic features and of the same theories which exhibit those features, will further be supported in chapter 10 to which historical material is devolved. The challenge for this section is to construct a model of the origin of aesthetic evaluative canons able to explain their variability, and the consequent variability of the judgements passed by scientific communities on the aesthetic features of given theories.

The mechanism to which the present treatment attributes the construction of an aesthetic evaluative canon operates as follows. A community selects its aesthetic canon at a certain date from amongst the aesthetic features of all past theories by attributing to each feature a degree of favour proportional to the degree of empirical success scored to that date by the set of theories which have appeared to embody that feature. The community's aesthetic canon is then composed of the set of mutually consistent such features which have won the greatest favour. A theory will consequently win support _ceteris paribus_ in the measure to which it shares the aesthetic features of past theories which have been attributed great empirical success.

Now, the degree of empirical success which the community attributes to long-established theories will in time decrease, either because of the discovery of fresh data unfavourable to them or because more recent theories will by comparison score greater success. Then the aesthetic features exhibited by the long-established theories will win a progressively lessening favour in the aesthetic evaluative canon, and a new theory's possession of those features will bring it less favour in the eyes of those who apply the canon in an instance of theory-choice. Simultaneously the empirical success of the recently-formulated theories will cause the aesthetic features of those theories to win greater favour in the canon, and any new theory to come before the canon will be valued for its possession of those new features. Thus evolve in time both the relative favour attributed to different aesthetic criteria in the canon and the preferences which the
canon expresses in cases of theory-choice.9

The degree of empirical success attributed to the set of theories exhibiting a given aesthetic feature may be altered by certain more specific considerations: a community may implicitly decide, for instance, to attach greater importance to an instance of predictive success gained in the immediate past rather than decades previously, or to successes gained in fields of active and vital research rather than those which are apparently stagnant. These preferences will exercise an effect on the favour which the aesthetic feature embodied by those theories will receive within the aesthetic canon.

In consequence of the mechanism outlined here, a theory which achieves significant empirical success will remodel its community’s aesthetic canon of theory-evaluation to attribute greater favour to its own aesthetic features. Thus the achievement of significant empirical success by a theory which exhibits given aesthetic features exercises an effect on subsequent theory-choice, since the aesthetic canon of theory-evaluation will as a consequence of that theory’s empirical success receive *ceteris paribus* a bias towards future theories which exhibit that theory’s same aesthetic features. By suitably altering the aesthetic canon to value its own aesthetic features, an empirically-successful theory will tend to replicate its own aesthetic features in the set of future theories which the community will come to embrace. A theory’s achieving signal empirical success at one moment in the history of science can thus influence the later development of science, or more precisely can help determine the set of theories which at later times of the history of science are embraced by the community in part consequence of its aesthetic canon of theory-choice.

The procedure described in the above paragraphs for the construction of an aesthetic canon is clearly a form of inductive inference, in that by its means the community is able to choose for its aesthetic canon the evaluative criteria which, under the stipulated definition of ‘fruitful’, have demonstrated the most fruitful applicability in the past. This procedure is however more aptly termed a metainduction rather than an induction in view of the level of its operation: while the term ‘induction’ customarily denotes a procedure which formulates or tests theories from an examination of past observations, this procedure formulates or

---

9 The mechanism outlined here for the construction of aesthetic canons of theory-assessment is similar to the mechanism of method-Darwinism by which Rescher (1977, pp. 140-66) suggests that methodological precepts are constructed. However Rescher believes that the precepts thus constructed are predictive of empirical success, while the present treatment holds that the aesthetic evaluative criteria constructed by the metainduction are not so linked.
tests methodological precepts from an examination of the empirical performance of past theories. For brevity the remainder of this treatment will refer to the procedure described in the previous paragraphs as ‘the aesthetic metainduction’.

The treatment which follows will be abbreviated by the adoption of two symbols. The function $x_i(t)$ will model the degree of empirical success attributed by the community to the set of theories which embody aesthetic feature $i$; $x_i$ has to be a function of time as both the discovery of new data and the invention of new theories which embody feature $i$ will cause the degree of empirical success attributed to the relevant set of theories to be altered. The function $y_j(t)$ will model the degree of favour which feature $i$ wins within the resulting aesthetic canon. The index $i$ ranges over the set of all possible aesthetic features of theories, such as all possible forms of simplicity, of symmetry, and so on; there is therefore in principle a separate application of each function for each distinguishable aesthetic feature to which one might refer in discussing the aesthetic appeal of a theory. In this metaphor, the operation of the aesthetic metainduction at time $t_0$ consists in the computation of the value of $y_j(t)$ for $t=t_0$ from a knowledge of the behaviour of $x_i(t)$ for $t \leq t_0$. Although it is probably fatuous to speculate at length on the functional by which one imagines the computation to be carried out, and which is therefore held to model the process of aesthetic metainduction, a suggestion of its form will be given in the following section once some characteristics of the metainductive procedure have been elucidated.\(^{10}\)

Fluctuations in the value of $x_i$ are meant in the present treatment to resemble, say, variations in the ratings of players of certain competitive sports such as professional tennis players. Victories or defeats in tournaments will increase or decrease the current rating of a player and his or her ranking in the sport; a victory in a Grand Slam final may furthermore be weighted more heavily, and thus have a greater repercussion in the ratings, than one in a provincial tournament. The resulting numerical rating given to a tennis player is thus an indication of current track record. Similarly, empirical success or failure will increase or decrease the value of $x_i$ currently attributed to a set of theories; furthermore an instance of empirical success in a test regarded as crucial may be weighted more in the eyes of the community, and thus exercise a greater effect on the value of $x_i$, than an instance of success in a less significant experiment. Thus the current value of $x_i$ represents the current track record of the set of theories exhibiting aesthetic feature $i$.

\(^{10}\) The relation between $y_j$ and $x_i$ is a ‘functional’ rather than a ‘function’ because both \textit{relata} are themselves functions, functions of time.
It is intended that the mechanism outlined here, and the claim that scientific communities use this mechanism to construct aesthetic canons of theory-assessment, be judged on two grounds: their ability to account for features of the history of science and the philosophical interest or fertility of the model of scientific rationality of which they form a part. These tests will be applied in later chapters, once certain consequences of the mode of operation of the metainduction have been laid out.

4. Aspects of the evolution of aesthetic canons

Let us consider the implications of the temporal variation of $x_i$ – the quantity representing the degree of empirical success enjoyed by theories exhibiting aesthetic feature $i$ – upon the construction of the aesthetic canon. The computation of the value of $y_i(t_0)$ from a knowledge of $x_i(t)$ for $t < t_0$ is a complex social procedure. The application of the aesthetic metainduction which is initiated by the formulation of a new theory incorporating aesthetic feature $i$ consists typically of the following steps: the diffusion of knowledge of the theory through the community; the estimation by experiment of the degree of its empirical success, a process corresponding in the metaphor to the determination of a value of $x_i$ which takes cognizance of the empirical success of the newly-proposed theory; the ascertainment of its aesthetic features $i$; and finally the computation of the new current value of $y_i$ from the perceived behaviour of $x_i$. The steps of this procedure each involve intricate social judgement and mediation, and will be apportioned among diverse institutions of the community including research laboratories, conference audiences and editorial boards. One can thus not expect that given a change in the value of $x_i$ the corresponding value of $y_i$ will be computed instantly, or that the aesthetic metainduction will be able to respond instantly to either the formulation of a new theory embodying feature $i$ or the discovery of data which causes a reassessment of the degree of empirical success to be attributed to such theories. On the contrary, the nature of the social network to which its operation is entrusted ensures that a damping will affect the aesthetic metainduction, and the product of the metainduction will react sluggishly to changes in the data or ‘input’ on which the metainduction is set to operate. In terms of the relational metaphor, there will be a time-lag between a change in $x_i$
and a corresponding change in $y$. In particular, the metainduction and consequently the aesthetic evaluative canon which it constructs will exhibit a delayed response to any unexpected change in the quality of the empirical performance of a theory: the unexpected success of the predictions of a low-rated theory or the sudden failure of a previously apparently reliable theory will be reflected only after some delay in the aesthetic canon of the community.

The time-lag which affects the functioning of the aesthetic metainduction has two consequences which are important for the model of scientific rationality under construction here.

The first consequence derives from the following observation. It has been established above that changes in the community's aesthetic evaluative canon lag behind changes in the community's perception of the degree of empirical success of theories. But the community's perception of theories' empirical success is the product of the application of the community's canon of logico-empirical criteria of theory-evaluation. Thus the time-lag of which the previous paragraph spoke may be seen as intervening between theory-evaluations or theory-choices on logico-empirical grounds on the one hand and those on aesthetic grounds on the other: theory-choices performed on aesthetic criteria will therefore lag behind theory-choices on empirical criteria, in that the recommendations for theory-choice issued by aesthetic criteria at a certain time will tend to resemble the recommendations which would have been yielded by empirical criteria at an earlier time in the history of science. Aesthetic judgements of theories in a community will tend forever to appear conservative or retrograde by comparison to empirical evaluations.

T.H. Huxley's aphorism about 'the great tragedy of Science — the slaying of a beautiful hypothesis by an ugly fact — which is so constantly being enacted under the eyes of philosophers', although probably penned as a rueful comment on the emotional commitment of scientists to their theories, describes one of the visible consequences of the lag of aesthetic appreciation behind empirical assessment. The perceived beauty of the hypothesis is a reflection of the accord between its aesthetic features and the preferences of the community's aesthetic evaluative canon, which as a consequence of the canon's metainductive origin lag behind the preferences which would be expressed by application of the logico-empirical evaluative canon; the aesthetic canon has not as yet reached the state of development which would enable it to value the aesthetic features of the

---

11 Huxley (1894), p. 244.
empirically more successful theory, or ‘fact’, which thus appears aesthetically unattractive or ugly. In time the community’s aesthetic canon will evolve by metainduction to consider more favourably the aesthetic features of the empirically-successful theory.

An aesthetically innovative theory will thus in all likelihood be seen as perceptually unappealing for a period after its first airing, since its aesthetic features neither conform to the pre-existing evaluative canons nor have yet had the opportunity to alter the canon to their favour; on the contrary the sign of a long-established theory is its being considered aesthetically pleasing by its community, an achievement due to the fact that its long and successful empirical track-record has ensured that its aesthetic features are now accorded favourable representation in the community’s aesthetic canon. An aesthetically original theory which in time achieves significant aesthetic success will therefore tend to follow the life-cycle of the bird dismissed as an ugly duckling when juvenile and acclaimed as a beautiful swan in maturity: while at first, in its novelty, the new theory appears aesthetically less pleasing than other theories of longer-established aesthetic form, its empirical success will gradually increase its aesthetic appeal. The theory’s aesthetic progression differs from that of the cygnet in one fundamental respect: the community’s aesthetic approval is won by the bird by changing its morphology to conform to the existing aesthetic canons of evaluation applied to members of its species, and by the theory by redefining the aesthetic canons applicable to it in accordance with its own aesthetic features.

Previous paragraphs introduced the variables $x_i$ and $y_i$ as elements of a possible mathematical representation of the procedure of the aesthetic metainduction. Since there is probably no hope of ever associating quantitative values to these parameters, great effort directed at proposing a precise functional which should be held to model the aesthetic metainduction would be misplaced. Nonetheless, the features of the metainduction described above bring to mind a particular such functional. An integral of a function over time gives an indication of the past behaviour of that function, being – roughly – the cumulative measure of the values which the function attained in the period under consideration. An integral of $x_i(t)$ would thus exhibit the behaviour required here of $y_i(t)$ in retaining a memory of the past behaviour of $x_i(t)$. Furthermore, the existence of a time-lag between changes in time of the variable $x_i$ and corresponding changes $y_i$ supports the idea that the relation between the two variables will be similar to an integral. Integration is well suited to model the damped response of a variable to changes in time of another; the behaviour of electronic devices which damp changes in
current, for instance, is appropriately modelled by integration. The relation which one may envisage between the degree of empirical success of theories exhibiting aesthetic feature $i$ and the favour attributed to that feature in the aesthetic evaluative canon is therefore of the form $y_i(t) = \int x_i(t) \, dt$. This functional would capture some of the characteristics of the aesthetic metainduction, most notably the sluggish response of changes of $y_i$ to corresponding changes of $x_i$.

While the sight of this relation may aid the mathematically-adept reader to visualize the operation of the metainduction, no further effort will in the present treatment be made to mathematicize the scientific community’s construction of an aesthetic evaluative canon.

Since scarcely any works in the philosophy of science devote committed attention to aesthetic criteria of theory-assessment, the number of those which understand the significance of the time-lag of aesthetic appreciation behind empirical judgement is virtually nil. A couple of authors have however expressed insights which accord with the above treatment. J. Bernstein has written:

> In science as in the arts, sound aesthetic judgements are usually arrived at only in retrospect. A really new art form or scientific idea is almost certain at first to appear ugly. The obviously beautiful, in both science and the arts, is more often than not an extension of the familiar. It is sometimes only with the passage of time that a really new idea begins to seem beautiful.  

The present treatment specifies a mechanism to explain what remains for Bernstein merely a phenomenological observation: the ‘passage of time’ which ensures that a novel, empirically-successful theory will appear beautiful is required for the social operation of the aesthetic metainduction. The view of Penrose is equally consistent with the remarks above:

> Perhaps one’s aesthetic judgements will change […]. Such judgements are, in any case, often to a considerable extent, acquired tastes. In these cases one cannot really appreciate the beauty of something until some familiarity with it has been gained – one has really to have thought about it for quite a while.

Of course for a new theory to win the approval of an aesthetic canon of a community it is not sufficient to ‘think about it’ or mull it over: it is necessary for the community to alter its aesthetic canon in response to the empirical success of the new theory and in accordance with the new theory’s aesthetic features.

A phenomenon which is distinct from the occurrence of a time-lag in the

---

12 J. Bernstein (1979), p. 3.
CHAPTER SEVEN: THE INDUCTIVE CONSTRUCTION OF AESTHETIC PREFERENCE

aesthetic metainduction but which nonetheless resonates strikingly with it was perceived by B. Russell. His concern in the following passage is for not substantive methodological beliefs in science but the literary forms or styles in which philosophers choose to couch their ideas. He believes that conventional or conservative philosophical doctrines will appear more beguiling than new ones through being expressed in well-worn prose:

Broadly speaking, old conceptions have acquired pleasant literary clothes, whereas new ones still appear uncouth. An aesthetic bias in favour of good literary form is therefore likely to be associated with conservatism. [...] As a result of many centuries of Platonism, the language of educated men can now express even the most difficult of Plato’s ideas without crabiness; but this was not the case in his own day. [...] In such ways those who insist upon elegant literary form are compelled to lag behind – often far behind – the best thought of their time. Per contra, conservatives have a great aesthetic advantage over innovators, for ideas [...] grow more beautiful as they grow older.14

The parallelism between Russell’s remarks on the prose style of philosophical works and the present treatment of aesthetic canons of theory-assessment is clear: in both cases the aesthetic garb of the substance is the more appealing for having been incorporated into the canons of taste of the community.

Naturally, the entrenchment of certain aesthetic commitments into the evaluative canons of a community is a phenomenon frequently witnessed in literary disciplines. As a literary critic, Nicolson has remarked upon the resilience of certain aesthetic devices in English poetry. Her interest is attracted by the appeal to scientific or natural-philosophical imagery in literary work, and her remarks on the time-lag suggest similarities between the behaviour of aesthetic canons in literature and that of those in scientific methodology:

I shall try to explain [...] how and why the Circle of Perfection finally broke under the impact of seventeenth-century science, and at the same time suggest that old habits die hard, and that time-honored ways of thinking about the world and man did not change in a moment. The time-lag has sometimes been as clear in scientists as in poets.15

Nicolson’s explicit reference to a ‘time-lag’ between the empirical results brought to bear in a discipline and the aesthetic canons which dominate creativity in that discipline suggest the intriguing possibility that the same processes of evolution of conceptual categories may be at work in literature and in science.

---

14 B. Russell (1940), p. 457. Russell prefaces these remarks his praise for the literary qualities of the works of Santayana.
CHAPTER SEVEN: THE INDUCTIVE CONSTRUCTION OF AESTHETIC PREFERENCE

The second consequence of the time-lag is the following. The aesthetic metainduction was metaphorically represented above as the computation of values of \( y_i \) for values of \( x_i \) changing in time. The variable \( x_i \) represents the degree of empirical success currently attributed by the community to the set of theories which embody aesthetic feature \( i \). There is clearly in principle no upper limit to the rate at which the value of \( x_i \) can change in time. Its value may increase quickly if, for instance, a newly-formulated theory embodying aesthetic feature \( i \) suddenly solves an outstanding empirical problem of its discipline; conversely the value of \( x_i \) may drop sharply if a previously well-regarded theory exhibiting feature \( i \) suffers a catastrophic empirical failure. High rates of change of \( x_i \) may however cause difficulties in the metainduction which aims to compute corresponding values of \( y_i \). It has already been conjectured that changes in \( x_i \) are reflected in changes in \( y_i \) only after a certain time-lag because of the sluggish capacity of response of the social institutions delegated to perform the metainduction. It is then a plausible further supposition that if the rate of change of \( x_i \) attains excessively high values the metainduction will begin to lose ground, and the time-lag which elapses between a certain change of \( x_i \) and the corresponding change of \( y_i \) will increase. Even when the time-lag between changes in \( x_i \) and corresponding changes in \( y_i \) is minimal, that time-lag may be reflected in minor discrepancies between the theory-preferences expressed by logico-empirical and by aesthetic criteria: these discrepancies arise from the fact that, because of the time-lag introduced by the metainduction, the aesthetic evaluative canon has a conservative bias in theory-choice of which the logico-empirical canon is free. When however the time-lag lengthens, the discrepancies between the preferences expressed by logico-empirical and aesthetic canons will worsen. When thus the rate of change of \( x_i \) and hence the length of the time-lag attains great magnitude, the practice of theory-choice in the community will lose the determinacy which in normal times is assured by the metainductive co-ordination of the preferences yielded by the aesthetic and empirical canons. High rates of change of \( x_i \) will thus render theory-choice controversial, and raise the prospect that scientists in the community may be compelled to elect between conducting theory-choice on empirical grounds and maintaining their previous commitment to the community’s aesthetic canon.

These observations will be exploited in the account of the advent of
revolutionary crises in science to ensue in chapter 8.\textsuperscript{16}

5. The psychological motivation of the metainduction

According to the present account of the origin of aesthetic canons of theory-assessment, aesthetic evaluations of theories are in general of no use in attempts to identify the empirically most successful theory among several which may be proposed. This incapacity of the aesthetic canon derives in principle from the fact that aesthetic canons evaluate theories on criteria unrelated to empirical virtues, and in practice from the fact that the metainductive construction of aesthetic canons (or, to employ the relational metaphor once more, the computation of values of $y$, from changing values of $x$) requires non-zero time. Because of this incapacity, scientific communities cannot legitimately turn to aesthetic canons of theory-evaluations to further the utilitarian or empirical interests which they repose in science and its progress.

The problem therefore arises of what rational reason, if any, should be attributed to scientific communities for constructing and applying aesthetic canons of theory-evaluation. The proposal of the present treatment is that there is no such rational reason, and that on the contrary the pursuit of the aesthetic metainduction is an arational phenomenon in scientific communities in the sense in which for D. Hume the pursuit of inductive inferences is an arational phenomenon in the mind of the individual.\textsuperscript{17} Thus, the present account of the origin of aesthetic canons of theory-assessment is explicitly intended to resemble

\textsuperscript{16} There is a degree of independence between two of the claims of this treatment which I should like to indicate. The validity of the model developed in the above two sections of this chapter to account for the community's construction of a certain canon of theory-assessment is independent of the validity of the claim that the canons of theory-assessment thereby constructed are aesthetic in nature. The mechanism of the aesthetic metainduction may well operate in scientific communities for the construction of certain components of the overall canon of theory-assessment even if it is denied that communities make recourse to aesthetic evaluative criteria for their assessment of theories. Needless to say there are other reasons, adduced in chapters 3 and 4, for which I should be reluctant to withdraw from the full complexity of my view, but the endorsement of the greater part of this chapter does not require the acceptance of the chapters that have preceded it.

\textsuperscript{17} I think of Hume (1739), pp. 69-179. The interpretation of Hume's Treatise as at least in part an essay in the natural science of the mind seeking to account for inductive inferences as a psychological phenomenon - upon which the analogy outlined in this section is predicated - is advanced in e.g. Stroud (1977), pp. 42-95.
the Humean explanation of the origin of notions of cause: just as Hume believed
the inductive apprehension of causal links to be unsupportable by nomological
data but a nonetheless ineluctable product of a driven mind, aesthetic canons in
science can boast no systematic relation to empirical virtues but spring from the
psychological concerns of scientists.

The psychological motivations of the Humean induction and of the
aesthetic metainduction are similar. Hume envisaged that the individual would
pursue inductive inferences out of a tendency to conceive that law-like
regularities held between events called ‘causes’ and other events called ‘effects’;
the present treatment similarly envisages that a scientific community pursues the
aesthetic metainduction out of a tendency to conceive that law-like regularities
hold between a theory’s achieving empirical success and its possession of certain
aesthetic features. The agent envisaged by Hume’s theory, recalling a previous
occasion on which the sight of flames was accompanied by the sensation of pain,
is moved to recoil from the location at which flames are again sighted: similarly,
the scientists contemplated by the present model of scientific rationality, recalling
a previous occasion on which a theory’s possession of certain aesthetic features
was accompanied by that theory’s achievement of empirical success, are moved
on subsequent occasions to value and pursue other theories exhibiting the same
aesthetic features.

In each case an inductive and thus – as Hume would put it – arational
expectation is conceived: Hume’s agent expects arationally that the renewed
proximity to the flames will bring pain, while the scientists of which the present
treatment speaks believe arationally that the possession by new theories of the old
aesthetic features will bring renewed empirical success. In each case, however,
analysis reveals the illusory status of the supposedly nomological correlations
invoked by the protagonists of the episode. Hume’s analysis concluded that the
notion that there could exist law-like correlations between perceived causes and
effects was philosophically nonsensical; similarly the analysis contained in the
above treatment concludes that there exist no law-like correlations between a
theory’s empirical success and its possession of given aesthetic features. In both
this account and that of Hume, the belief in such correlations cannot be the result
of an inference from empirical observations but is rather a psychological artefact.

It is important to acknowledge that neither Hume’s account nor this one
concludes that correlations which are grasped by a process of arational induction,
and which therefore cannot claim an objectivist derivation, are for this reason
valueless. Links of cause and effect are an undeniable convenience of the Humean
life, and Hume states that appeal to such links in the planning and performance of quotidian actions is helpful to their success and indeed to the preservation of the life of the agent. Similarly the construction – by means of the aesthetic metainduction – of correlations between theories’ empirical success and their possession of certain aesthetic features is often an aid to the processes of scientific decision-making: in phases of science when the value of $y_i$ is reacting promptly to changes in $x_j$, the belief in such correlations reinforces the community’s tendency to make the empirically most fruitful choices among competing theories.

The pragmatic utility of correlations constructed by Humean inductions naturally does not absolve the observer of the process from the obligation to remember their subjective nature and to avoid attributing to them any necessity.
1. Gradualism and catastrophism in the philosophy of science

The foregoing treatment of aesthetic criteria of theory-assessment is in principle consistent with the thesis that canons of rationality are subject to only continuous change. The discussion of the two categories of criteria of theory-assessment in the previous chapter attributed to logico-empirical criteria stability through time and interpreted aesthetic canons as undergoing a continuous evolution in consequence of the process of aesthetic metainduction. These attributions leave open within the model the possibility that there will be no discontinuous changes in the values pursued by scientists in theory-choice, and hence that canons of scientific rationality will evolve only gradually and continuously. If this model of rationality were instantiated in scientific practice, then although the effects of the small gradual changes might summate to yield in time different rational canons, any reasonably short interval of the history of science would exhibit a continuity of methodology and certainly contain no discontinuous substitutions of one methodological canon by another.

The project of reconstructing scientific rationality would probably regard this conclusion as eminently satisfactory were it not for the spur of historiography. Chapter 2 made reference to the power of historiographic findings to undermine the conviction that logico-empirical criteria are sufficient to reconstruct scientists' practice of theory-choice: similar findings now shake the belief that canons of rationality are immune from discontinuous alteration in the history of science.¹ The persuasion that the evaluative canons of scientific communities must undergo discontinuous changes acquires force from, for

¹ The influence of the findings of history in persuading philosophical observers that scientific progress suffers methodological discontinuities is proclaimed by Kuhn (1962, pp. 1-9) and studied by I.B. Cohen (1985, pp. 389-404). Notes on which categories of historiographic findings are most apt to suggest to the historian the occurrence of a revolution may be found *ibid.*, pp. 40-7.
instance, a study of the nature of the arguments advanced in theory-assessment in early-modern Western science. Copernicus adduces among the grounds for the heliocentric model of the solar system the observation that, if this model constituted an accurate representation of the solar system, an analogy could then be drawn between the configuration of celestial bodies and the organization of a monarch's court. The analogy postulates that the pre-eminent being of each structure occupy its centre and be surrounded by dependants or satellites: 'Thus indeed the Sun as if seated on a royal throne governs his household of Stars as they circle round him.' For this analogy to possess the advocative force with which Copernicus implicitly invests it, Copernicus's methodological canon - or, more realistically, that enshrined in his community - must have included a precept to the effect that a theory is *ceteris paribus* to be preferred if the mechanisms which it posits admit relations of analogy with contemporary social or political institutions. The philosopher appraised of the apparent adherence to this precept in sixteenth-century communities of planetary astronomers is likely to postulate the occurrence of a discontinuous switch in methodological canons in that community between the time of Copernicus and, say, that of Newton. Surely, the philosopher wonders, no process of merely inductively reviewing methodological precepts could fashion modern mathematical science from medieval scholastic speculation?

If, on the one hand, in order to account for historiographic findings it appears necessary to embrace the catastrophist supposition that moments of revolution have witnessed considerable sudden changes of rational canon, on the other hand two gradualist conclusions emerge equally compelling from the historical record. The first is that revolutions are separated from one another by periods in which theory-succession is governed by an unchanging methodological canon. The second is that even in revolution not every norm of scientific practice can be abrogated and replaced by a different precept: if revolution visited such a thorough methodological metamorphosis upon a community, there would presumably remain no sense of a unified history of science or indeed of a historically identifiable scientific discipline.

---


3 The historian's reaction to the discovery that Copernicus appealed to arguments of this kind is examined in Hesse (1973), p. 137, where it is suggested that the historiographic practice of assessing the truth and standards of rationality of past scientific theories may help to discern the occurrence of revolutions in methodological canons.
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

In view of these historically-founded apprehensions, the philosophy of science faces the task of accommodating within its model of the historical development of science both elements of continuity and elements of discontinuity. It would be a considerable achievement for any theory of scientific rationality to combine these elements satisfactorily, seeing in science a thread of procedural continuity whilst explaining how at times of revolution that thread can be frayed and thinned in a partial change of methodological practice.

The acknowledgement, under the pressure of historiographic findings, that continuous scientific progress is punctuated by discontinuous changes in evaluative canons has occupied the philosophy of science since at least the 1930s when G. Bachelard wrote of *ruptures épistémologiques* and L. Fleck of the alternation of successive *Denkstile*. The most influential account of discontinuous changes of rational canons is however today undoubtedly that of Kuhn.

He explicitly recognizes the need for models of science to accommodate both phenomena of continuous change embedded in tradition and phenomena of discontinuous innovation, between which he believes there endures an ‘essential tension’. Kuhn views intellectual history as sectioned into periods of ‘normal science’ defined by paradigms during the life-span of each of which the progress of science adheres to unchanging methodological canons. These periods are terminated by instances of ‘radical standard variance’, or ‘changes in the standards governing permissible problems, concepts, and explanations’.

Some difficulty has however been encountered in fathoming from Kuhn’s writings in what this variance is supposed to consist, or which specific categories of norms of scientific procedure are liable to such discontinuous changes. Kuhn has volunteered no adequate elucidation of this matter. On the contrary, the same essay of his which introduced the prospect of radical standard variance proceeded to offer a ‘a preliminary codification of good reasons for theory choice’, a tabulation of criteria of theory-evaluation to which Kuhn apparently attributes a

---

6 Kuhn (1962), p. 106. Successive paradigms are in Kuhn’s original view separated also by a ‘radical meaning variance’ of the terms which they employ, but this variance would not of itself endanger the continuity of canons of rationality and hence will not be addressed here. In any event Kuhn has withdrawn from this extreme position to allow for the communication of meaning between proponents of different paradigms: see *ibid.*, pp. 198-9.
7 This phrase occurs in Kuhn (1970), p. 261, whereas the passage which this phrase is intended to describe is contained in his (1962), pp. 144-55.
validity which is ever-lasting or at least endures through paradigm-changes. Clearly the identification of such ‘good reasons’ would much reduce the prospect of radical changes of criteria of theory-evaluation in science.

Kuhn lent further weight to the gradualist component of his model in a later work. He there gave explicit form to the promised codification of science’s methodological precepts, professing the opinion that there exist five criteria for theory-evaluation which will be common to the proponents of all paradigms: the qualities of accuracy, consistency, breadth of scope and fruitfulness. Kuhn stresses that ‘it is vitally important that scientists be taught to value these characteristics’: the historically indiscriminate reference to ‘scientists’ as if their methodological canons were not bound to some particular paradigm and as if their theory-choices were throughout history to be based on Kuhn’s five values tends further to vanify Kuhn’s painstaking earlier work in allowing for the possibility of radical variance in their canons. As long as the values of theory-choice applied by scientists are reputed to possess trans-paradigmatic validity and no complementary class of paradigm-specific norm is identified, it remains unclear how radical standard variance could ever arise.

From these observations it appears that Kuhn’s treatment of criteria of theory-evaluation, taken in its entirety, is unable satisfactorily to mediate between the radical desire to allow for sharp discontinuities in scientific procedure and the moderate realization that some communication and sharing of concerns and preferences across such purported discontinuities is revealed in history. Instead of combining elements or mechanisms of continuous development and of discontinuous substitution within a unified picture of the progression of science, Kuhn’s writings appear to oscillate uneasily between postulations of outrageous catastrophism and of tame gradualism.

The root of Kuhn’s incapacity to achieve the desired synthesis lies, it is here claimed, in his treatment of all norms of scientific methodology as exhibiting the same historical nature and behaviour. Kuhn does not draw internal distinctions within his repertory of five values of theory-assessment: he does not...

---

10 Kuhn (1970, p. 262, and 1977, pp. 322-5) attempts to allow for radical standard variance in extremis by supposing that although the parties to a dispute may agree on the criteria of theory-evaluation, they could still differ in the application or the relative weighting of such principles in certain cases of theory-choice: but this is patently a case not of a radical variance of standards, merely of their indefiniteness or of the subjectivity of their application.
11 I think of Kuhn (1962), pp. 111-35, as an example of the former tendency; of his (1977), pp. 31-65, as an example of the latter.
envisage the possibility that some of these values may possess a degree or form of historical variability different from that of certain others on the list. A distinction of this kind would allow Kuhn to adduce some criteria as liable to radical variance and thus as responsible for revolutionary discontinuities in theory-evaluation, whilst leaving to other criteria the trans-paradigmatic validity which he is apparently keen to see in his revised treatment of scientific revolutions. A view of criteria of theory-assessment which differentiated amongst them along these lines would also possess considerable historical verisimilitude, since it could accommodate the views of both metainductivists who perceive in evaluational criteria historical variability and metarationalists who maintain that at least some such criteria have remained unchanged: these apparently incompatible tenets would be seen each to capture part of the complex behaviour of methodological norms in history.

2. Aesthetic canons in normal science

The reconciliation of elements of continuity and of discontinuity within the same model of scientific rationality will be attempted in this chapter by appeal to the considerations on aesthetic criteria of theory-evaluation laid out so far in this work. The attempt will proceed, as anticipated in the previous section, by postulating that some but only some of a community’s methodological criteria are subject to the kind of radical variance which is envisaged by Kuhn.

The present view of criteria of theory-assessment is well able to erect a differentiated model of their operation in scientific revolutions, for it has already and on independent grounds distinguished two classes of methodological norms which differ in historical origin and behaviour. Logico-empirical evaluative criteria were above attributed a metarationalist origin and temporal permanence, while aesthetic criteria were reputed to be the product of a continuous metainductive updating. In this section and the next, a model of scientific revolutions will be constructed which delegates one of these classes of evaluative norms to assure the elements of continuity in scientific methodology across revolutions, and details the other class to account for the elements of revolutionary discontinuity, embodying in their changes the differences between pre- and post-revolutionary procedural canons. More specifically, and – by this
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

point – obviously, the present view attributes the elements of historical continuity in scientific procedure to the permanence of logico-empirical evaluative criteria, and allocates responsibility for the occurrence of revolutionary changes of methodology to substitutions of one canon of aesthetic evaluative criteria by another. On this view a revolution will consist of a discontinuous change not in the criteria of theory-evaluation in their globality, as Kuhn appears to believe by default of any more differentiated statement, but solely in the criteria of aesthetic evaluation endorsed by a community. The pre- and post-revolutionary states of a science will be united by their empirical criteria in consequence of the a priori formulation of the latter, but will be distinguished by their appeal to two different sets of aesthetic criteria.

This view will be presented here partly by reference to the terms and concepts coined by Kuhn’s model of scientific revolutions: this use of an already well-known conceptual framework will ease the exposition, and in time will render obvious the differences between the present view and that of Kuhn.

The previous chapter outlined a model of the development of an aesthetic canon of theory-appraisal by a scientific community. This development is continuous, in that the formulation of the canon is continuously reviewed and updated by the procedure of aesthetic metainduction defined in that chapter. A community constructs its aesthetic canon at a certain date from among the aesthetic features of all past theories by attributing to each feature a degree of favour proportional to the degree of empirical success scored up to that date by the theories which have appeared to embody it. The collection of aesthetic features thus constructed assumes a normative role, becoming a canon of theory-assessment: a newly-formulated theory will be evaluated by the community partly for its accord with the aesthetic canon thus constructed and partly for the degree to which it satisfies the logico-empirical desiderata separately imposed upon theories by the community.

It was further explained in the previous chapter how this procedure ensured that aesthetic evaluative canons were tendentially conservative: they will forever tend to prefer, and to recommend for preference by the community, theories which are ‘more of the same’ in conforming to the aesthetic features embodied by the empirically most successful theories of the recent past. In some instances of theory-choice the in-built conservatism of the aesthetic evaluative canon may dissuade the community from readily embracing the empirically most successful theory, if that theory faces a competitor which is aesthetically more acceptable in better conforming to the community’s aesthetic canon: such a case,
which holds great significance for the consummation of a phase of scientific revolutions, will further be explored below.

There equally exists a possible period of scientific history, which it is best to consider first, in which on the contrary the conservatism of the aesthetic canon exerts no undue brake on the logico-empirical progress of science, or more precisely in which the aesthetic canon to no extent weighs against or hinders the scientific community's adoption of the scientific theories which are at each moment exhibiting the greatest empirical success. This state of affairs persists while there remains, as a matter of contingent fact, agreement between the theory-choices recommended by application of the logico-empirical criteria and those recommended by application of the aesthetic canon of the community. This requirement may be more concisely expressed by once more representing the aesthetic metainduction as the social computation of values of \( y_i \) (the degree of favour attributed to aesthetic feature \( i \) in the community's aesthetic canon) from changing values of \( x_i \) (the degree of empirical success attributed by the community to the set of theories which embody aesthetic feature \( i \)). As was shown in the previous chapter, a community's logico-empirical and aesthetic canons of theory-evaluation will maintain reasonable agreement in cases of theory-choice as long as the time-lag between changes in \( x_i \) and corresponding changes in the values of \( y_i \) computed by the process of aesthetic metainduction remains reasonably short. This result is due to the fact that, under such conditions, changes in the community's judgement of the empirical success of their theories will be reflected reasonably promptly in the preferences expressed among the same theories by the community's aesthetic canon. Whilst the time-lag remains reasonably short, theory-choices performed on the community's aesthetic canon would not yield theories other than those which would be selected by the separate application of the community's logico-empirical criteria: the joint and at times not explicitly distinguished application to problems of theory-choice of both sets of criteria would naturally further reinforce the choices of each.

Thus, for this phase of scientific history to endure, the rate of evolution of a community's aesthetic canon must be sufficiently high to ensure that the time-lag mentioned above remains 'reasonably short', and therefore that theory-judgements yielded by the application of the aesthetic canon and those passed by application of the logico-empirical canon remain in reasonably close agreement. For the rate of evolution of the aesthetic canon to be 'sufficiently high' to ensure this result, it need not be especially high: the aesthetic canon's agreement with the logico-empirical canon is equally assured if, say, the aesthetic canon remains
unchanged during a period in which all empirically successful theories to be formulated are aesthetically conservative, i.e. satisfy the requirements of the extant aesthetic canon. The accord between a sequence of newly-formulated theories and the dominant aesthetic canon may in other terms be obtained equally by a combination of a static aesthetic canon and a low degree of aesthetic innovativeness in theory-formulation, or by a combination of high aesthetic innovativeness in the latter practice and a sufficiently high rate of evolution of the aesthetic canon: on the present account those two cases do not differ essentially.

Of course the agreement in theory-preference between the logico-empirical and the aesthetic canons breaks down as soon as the community encounters a scientific theory which a) violates the stipulations of its aesthetic canon, and b) exhibits empirical success superior to that of its competitor-theories which conform to the canon. At such a time one canon would recommend adoption of one theory and the other would express preference for its rival. Yet, because of the mechanism of the construction of the aesthetic canon, a persistence of the state of agreement between the two canons is likely to entrench itself in the community’s patterns of theory-choice. As long as there remains a coincidence between the aesthetic features prescribed by the canon and those embodied by the empirically most successful proposed theories, two mechanisms of positive feedback will manifest themselves. First, the empirical success of each newly-proposed theory will – under the operation of the aesthetic metainduction – further increase the degree of favour attached to the aesthetic features exhibited by those theories within the aesthetic evaluative canon: for its part, and in consequence of this first phenomenon, the aesthetic evaluative canon will attach increasing value to the same aesthetic features which happen to be exhibited by the empirically most successful theories proposed during the period. Thus, though the persistence of this period of agreement between the logico-empirical and the aesthetic evaluative canons depends on a continued coincidence, this coincidence is liable to entrench itself and the resulting acts of theory-preference in the behaviour of the community.

The phase of science which has thus far been described corresponds, it is here stipulated, to a period of ‘normal science’ in the Kuhnian model of scientific history. A period of normal science is on this view, as on Kuhn’s, one in which theory-choices are uncontroversial and acts of theory-choice generally attract the consensus of the entire community. This view attributes the consensual nature of

---

12 Kuhn’s characterization of ‘normal science’ is to be found in his (1962), pp. 23-34.
theory-choice in periods of normal science to the fact that in such periods theory-
choices based on empirical criteria and those founded on aesthetic criteria coincide and hence there does not arise the dilemma of choosing between the aesthetic appeal of some theories and the observational success of others. This view explains also how the consensus is entrenched in the community, by the mutual reinforcement of the choice of a sequence of aesthetically similar, empirically-successful theories on the one hand and the aesthetic evaluative canon on the other, in the manner described in the previous paragraph.

The explanation of the consensus achieved in normal science which is offered by the present treatment attains a level deeper than that invoked by Kuhn's model: whereas that model limits itself to postulating the occurrence of periods of community-wide satisfaction about the speed and manner of the progress of a branch of science, this view reconstructs the satisfaction as the rational recognition of a persisting agreement in cases of theory-choice between the logico-empirical and the aesthetic classes of evaluative criteria.

This characterization of periods of normal science permits historians of science to form certain expectations about the manner in which such periods will appear to them in the historical record. A period of normal science defined in the way suggested here will spawn a sequence of theories which displays unchanging or smoothly varying aesthetic features. The historian aiming to discern a period of normal science will thus search for a run of theories united by aesthetic resemblance. For instance, it will be argued in chapter 10 that the development of Western planetary astronomy in the interval from the work of Ptolemy to and including that of Copernicus constituted a period of normal science since the empirically successful theories formulated in that discipline and in that interval exhibit significant common aesthetic features, viz. principally the metaphysically-backed canon of simplicity or symmetry which attributed to celestial bodies circular paths.

The 'paradigm' which Kuhn sees as characterizing work within a period of normal science is on this account defined by the aesthetic canon which contributes to theory-choice within that period. Once this identification is made, a difference readily emerges between the content of Kuhn's notion of a paradigm and that of the corresponding notion promulgated here. Kuhn at least implicitly attributes to a paradigm a degree of rigidity or fixedness in time, and certainly his model of the progression of science omits to mention a mechanism by which a paradigm may evolve: in Kuhn's model, changes in canons of theory-preference are attributed entirely to revolutionary exchanges of one paradigm for another,
and the phases between revolutions are intended to witness only increases in the problem-solving capability of the set of theories which satisfy the evaluative stipulations of the paradigm then current. In the present model, on the contrary, a paradigm is attributed a mechanism by which to evolve during its own period of dominance of a community. The process of aesthetic metainduction by which – in the account of the previous chapter – an aesthetic canon of theory-evaluation is constructed does not terminate with the canon’s attainment of dominance in a community, but continues to update the canon in accordance with the community’s changing judgements of the degree of empirical success scored by its theories. Since the prescriptive content of a paradigm is in this treatment defined by the criteria enshrined in the aesthetic evaluative canon which characterizes that paradigm, it follows that the prescriptive content of a paradigm may undergo evolutionary change – owing to the operation of the aesthetic metainduction – even during the period of recognized dominance of that paradigm.

Kuhn terms ‘puzzles’ the staple problems of normal science which are solved within the precepts of the dominant paradigm.13 If now a paradigm is defined by an aesthetic evaluative canon, it follows that on this view puzzles must be found solutions constituted by theories or extensions of theories which accord with the aesthetic canon in force. On this view, as on Kuhn’s, while such solutions may be difficult to find, their acceptability is not generally a matter of controversy as it is the essence of such contributions that they should accord closely – if need be in a pedestrian, derivative manner – to the stipulations of the paradigm. Again, chapter 10 will depict the Copernican planetary theory as such a solution to a puzzle, which accorded deliberately with the requirements issued by the aesthetic canon then current.

Difficulties encountered by the community in attributing to ‘puzzles’ solutions which satisfy the dominant aesthetic canons may herald a crisis in normal science, as the next section will indicate.

13 Kuhn discusses the procedures of puzzle-solving ibid., pp. 35-42.
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

3. The abandonment of aesthetic commitments

As long as a community in a phase of normal science remains able to meet the problems thrown up in research with theories of which the choice is recommended by both its aesthetic canon and its logico-empirical criteria of theory-assessment, those problems retain the status of ‘puzzles’, the solutions which are attributed to them are accepted without controversy, and the mode of normal-scientific work continues uninterrupted. During such a period, the community entrenches the habit of demanding that its theories accord both with the logico-empirical criteria of theory-assessment and with the aesthetic evaluative canon which has dominated the period: the imposition of these two separate constraints on theory-acceptance poses no difficulty for theory-choice as long as the empirically most successful theories proposed within the community accord also with the dominant aesthetic canon.

In time however the community will encounter a kind of problem which is new and harder to solve.14 The difficulty will manifest itself in the following manner. Each problem of this new kind, just as any other problem, will attract a variety of proposed solutions; while however problems typical of the preceding period of normal science readily admitted solutions of which the adoption was recommended by both classes of evaluative criteria, for problems of the new kind the community’s two canons of criteria appear unable to agree upon which member of the class of proposed solutions to recommend. In the set of proposed solutions of a problem of the new kind, the one which accords best with the dominant aesthetic canon demonstrates less empirical success than do at least some of the proposed solutions which violate the canon.15

It appears that theory-choice cannot now satisfy the conjunction of the constraints separately imposed by the two evaluative canons. Out of the class of proposed solutions of a problem of the new kind, one will be recommended for adoption by the application of the logico-empirical criteria of assessment, and another will be preferred by the aesthetic evaluative canon. The previously enduring coincidence in acts of theory-choice between the recommendations of

14 As will become clear, problems of this kind correspond to Kuhnian ‘anomalies’, discussed by him ibid., pp. 66-76.
15 The sentence to which this note is appended is to be regarded as the definition of the expression ‘problem of the new kind.’ Thus, the present treatment envisages that a crisis of the type described in the remainder of this section will occur when the community encounters a problem which satisfies the stipulations of this sentence. The sentence which precedes the one to which this footnote is appended is on the contrary a characterization of the manner in which problems of this new kind will be recognized by the community.
the logico-empirical criteria of theory-assessment and those of the aesthetic evaluative canon is dissolved. The repercussions of this occurrence in the community are obvious: while the consensual nature of theory-choice in the period of normal science was assured by the agreement of the two classes of evaluative criteria, their conflict under the conditions described here will bring controversy to theory-choice.

The cause of this crisis can be retraced to the rate at which the community has found it possible to update its aesthetic evaluative canon by means of the metainduction. The previous section specified that the two evaluative canons of a community would continue, and therefore a period of normal science would persist, as long as changes in $x_i$ were reflected reasonably promptly by the metainduction in changes in $y_i$. A failure of the metainduction to maintain the lag short enough means that changes of the aesthetic canon will fail to keep pace with the evolution of the aesthetic features exhibited by the community's empirically most successful theories as these are formulated. As the time-lag between changes in $x_i$ and corresponding changes in $y_i$ increases, scientists will find that the aesthetic features of their empirically most successful theories increasingly diverge from the features required by the aesthetic evaluative canon, and thus that their empirically most successful theories come into deepening conflict with that canon.

When theory-choices based on the aesthetic canons of the paradigm have deviated from simultaneous judgements founded on empirical criteria so far that it becomes impossible for scientists involved in theory-choice to recompose or overlook their conflict, the scientists of the community are faced with the dilemma of choosing from between their evaluative canons the one of which to follow the directives. At first they may well deem that they can afford to suspend judgement on the issue, evading the embarrassment of adjudicating between the conflicting recommendations of their two classes of criteria for theory-choice. It is unlikely however that such an option will be considered legitimate indefinitely, and scientists will in time be forced to a choice between the conflicting evaluative canons and thence between the competing theories.

This is the moment of revolution. Under the conditions described, two factions of scientists will coalesce within the community, each of which will choose to resolve the crisis in theory-choice by relaxing one of the two sets of constraints imposed by the joint application of the two evaluative canons. This relaxation will in the case of each faction amount to suspending the application of one of the two evaluative canons, and conducting theory-choice exclusively by
application of the other canon.

Of the two factions of scientists, one will persist in adhering to the aesthetic norms characteristic of the paradigm in which they have intellectual commitment and will execute theory-choices under the guidance of the aesthetic canon alone, even though such a policy will generally cause them to pursue theories which are empirically less successful than those embraced by their rivals. The attitude of this group of scientists is captured by the pronouncement of H. Weyl: ‘My work always tried to unite the truth with the beautiful: but when I had to choose one or the other, I usually chose the beautiful.’ Since an aesthetic canon tends by reason of the time-lag introduced by the metainduction to be more conservative in theory-choice than is the logico-empirical evaluative canon, this group of scientists will here be termed the ‘conservative faction’.

The decision of this faction to suspend application of the logico-empirical canon and conduct theory-choice under the guidance of the aesthetic canon alone will be rejected by the other, more progressive party. Members of this party will make the opposite choice: they will suspend allegiance to the aesthetic canon, and conduct theory-choice on the exclusive basis of empirical criteria. Since this course of action relaxes the extra-empirical constraints on theory-choice, it will permit the progressive faction to adopt theories empirically more successful than those of their conservative adversaries.

The abandonment by a faction of scientists of the previously dominant aesthetic canon, and their conduct of theory-choice independently of aesthetic criteria in the exclusive pursuit of logico-empirical success, is the revolutionary act. In accordance with one’s intuitive expectations of revolutionary measures, this act consists largely of a disavowal of the erstwhile commitments of the community in which it is hatched: the progressive faction in the revolutionary phase perceives the previous aesthetic commitments of the community to have constrained theory-evaluation in such a way as to have hampered the achievement of the greatest logico-empirical success, and thus to have compromised the progress of science during its period of domination. The faction believes justifiably that the relaxation of those constraints on theory-choice will enable the community to perform the empirically most fruitful choices among competing theories, and thus to impress the greatest speed upon the growth of the empirical power of science.

The two factions of scientists may coexist in the community for a period.

---

However the difference between the degree of empirical success of the set of theories embraced by the progressive faction and that of the set of theories advocated by the conservative faction will continue to grow in favour of the former. One imagines that for some sufficiently great magnitude of this difference, the former set of theories will overcome the residual aesthetic resistance of conservative members of the community and impose itself as undeniably more attractive than the latter set. Scientists relax their commitment to the pre-existing aesthetic canons of their community in accepting to decide cases of theory-choice in favour of the theories which are empirically the most successful among those proposed, but which accord less well than do other proposed theories to the pre-existing aesthetic canons of the community. When the entire community has, as advocated by the progressive faction in its midst, relaxed its commitment to its pre-existing aesthetic canons of theory-assessment, the revolutionary phase is terminated.

The effect of the revolutionary interlude as described here has been to strip the scientific community of one of its two canons of criteria of theory-assessment: by the time of the completion of this phase, the community no longer recognizes a commitment to a canon of aesthetic criteria of theory-assessment and in its theory-choices is governed solely by the logico-empirical canon. Indeed, scientists working in the phase immediately after the abandonment in a revolutionary crisis of a previously dominant aesthetic canon may harbour resentment against the canon which they have abjured: they may see in it an influence which – at least in the latter stages of the period of its dominance – diverted the efforts of the scientific community from the pursuit of empirical success by imposing aesthetic and thus non-empirical commitments on theory-choice, hindering the empirical progress of science. They may in consequence nurture the positivistic hope that their community will dispense for evermore with aesthetic criteria of preference and conduct theory-evaluation with the sole aim of maximizing the empirical success of the community’s theories.

The present treatment suggests that their hope that science will from that moment onwards be guided by a methodological canon free of extra-empirical criteria of preference will be disappointed. After the end of the revolutionary phase which dislodged the aesthetic criteria from the community’s canon of theory-evaluation, the psychological concerns responsible for driving the aesthetic metainduction will return to affect the preferences of scientists. Scientists will begin arationally – and possibly unthinkingly – to postulate a correlation between the aesthetic features of the theories which have been adopted since the end of
the revolutionary phase and the empirical success demonstrated by those theories. They will, like true Humeans, come further to believe arationally that a new theory will be assured of empirical success by its possession of the aesthetic features associated with success in theories of the immediate past. Scientists will consequently hope to increase the empirical success of the set of theories to which their community lends assent by searching for those same aesthetic factors in new theories. To some of these features — feature i, say — will be associated great empirical success — that is, a high value of \( x_i \) — and scientists will tend to prefer theories exhibiting feature i in subsequent instances of theory-choice, in a procedure which amounts to increasing the value of \( y_i \) in accordance with increases in the value of \( x_i \).

By this psychological association and projection, the community recommences the operation of the aesthetic metainduction and undertakes the construction of a fresh canon of aesthetic criteria of theory-evaluation. A new phase of normal science is thus inaugurated, during which theory-choice will be conducted in part on the criteria of its characteristic aesthetic evaluative canon. In time, of course, divergences will begin to appear in cases of theory-choice between the preferences of the logico-empirical criteria and the new aesthetic canon, until a new revolutionary crisis and ultimately a further revolution precipitate.\(^{17}\)

4. The aesthetic interpretation of revolutions

The present model of scientific revolutions attempts to fulfil the requirement enunciated at the outset, that any such model should combine in the dynamic

\[^{17}\text{From the vantage-point now reached, the present treatment is able to comment on the hope nurtured by some methodologists and reported in chapter 5 that the introduction of aesthetic criteria among the grounds for theory-choice will assure the scientist of sufficient reason upon which to choose between any two theories, and even between a pair of empirically equivalent theories. The observation that the aesthetic evaluative canon which a community embraces is peculiar to a certain phase of normal science and is abandoned upon the advent of a revolution suggests that any decision on aesthetic grounds between empirically equivalent theories will in general be thought by the community to retain validity only for the duration on the period of normal science in which the decision was taken; once the aesthetic canon which reigned at the time of the decision is repudiated, the choice taken will in general appear unjustifiable and the need will be perceived for the new aesthetic canon to be applied to the case of theory-choice.}\]
which it attributes to revolutions both features of continuity and mechanisms of
discontinuous change. In this attempt, the present model lends weight to some of
Kuhn's insights into revolutions but disputes certain other of his affirmations.

This model offers an explication of Kuhn's intuition that whereas in normal
science there is widespread agreement within the community on what constitutes
solutions to the problems in hand, during revolutionary crises there is no similar
consensus among scientists over the principles of theory-choice that ought to be
applied. According to the present view, the former phenomenon is explained by
the accord which persists during a period of normal science between the two
canons of theory-assessment; the latter phenomenon ensues when one group in
the community following the directives of the empirical evaluative criteria while
another favours the aesthetic criteria then current.

However, the characteristic feature of the present model is its interpretation
of revolutions as discontinuous changes in a part and in only a part of a scientific
community's methodological canon. Whilst a revolution effects a discontinuous
substitution of one aesthetic evaluative canon by another, the community's canon
of logico-empirical evaluative criteria endures unaltered through the commotion.
This view therefore does not vindicate Kuhn's stronger claim that theories
embedded in rival paradigms cannot be compared since, as he believes, there are
no paradigm-neutral principles relative to which this comparison could be
executed: the trans-paradigmatic continuity of science is assured by its empirical
criteria even despite their repudiation by aesthetically conservative factions at
times of revolutionary crises. It is on the present model definitely untrue that 'the
normal-scientific tradition that emerges from a scientific revolution is not only
incompatible but often actually incommensurable with that which has gone
before': scientists of all paradigms share the values encapsulated by science's
empirical criteria.

A consequence of the interpretation of revolution as a change in only a
component of a community's methodological canon is the attribution of an
element of historical continuity to rationality. While scientists in different
paradigms reason differently in that they adhere to different canons by which to
perform the aesthetic evaluation of theories, they acknowledge the same canon of
rationality in as far as they share another set of criteria for theory-choice, the set
composed of the empirical evaluative criteria. Communication of reasons across
revolutionary discontinuities is therefore not wholly impossible: while aesthetic

---

18 Ibid., p. 103.
reasons for theory-choice would not be shared by the participants in such communication, each would acknowledge the other's concern for logico-empirical features of theories such as logical consistency or predictive accuracy.

This model is consequently able to offer a moderate interpretation of Kuhn's view that a revolution consists of a switch of Gestalt or of world view. A revolution, in as far as it consists of an exchange of aesthetic canons, and in as far as aesthetic canons determine a perceptual outlook, is a change of the expectations with which a scientific community perceives its scientific theories: to this extent Kuhn's view of revolutions as changes of perspective is justified. The interpretation of scientific revolutions as a change of aesthetic canons explains how J.H. Wheeler could think of extending to contemporary physics what Gertrude Stein said of modern art: 'It looks strange and it looks strange and it looks very strange; and then suddenly it doesn't look strange at all and you can't understand what made it look strange in the first place.' By 'a strange look' one readily intends, in science as in art, a violation of the dominant aesthetic canons of a time; a work of art or a theory cease to look strange when they convert the aesthetic aversion which they initially encounter in their community into aesthetic approval or acclaim. The procedure by which a theory converts its community to appreciation of its own aesthetic features captures some elements of an exchange of Gestalt; but since under this interpretation the effects of a paradigm-change do not implicate the community's logico-empirical evaluative canon, the resulting rupture is less perturbing of science's methodological canon than would be a genuine, all-pervasive, switch of world view.

Finally, the present treatment conflicts with the construal which Kuhn gives of revolutions, in at least his catastrophist writings, as irremediably irrational transitions governed by what his critic Lakatos has termed 'mob psychology.' Kuhn is led to this construal by his belief that, since there is no continuity of canons of rationality outside the temporal and conceptual bounds of a paradigm, there can be no 'rational reasons' for choice outside those bounds, and thus in particular there can exist no rational reasons on which to choose between competing paradigms. The present view on the contrary assures the existence outside the temporal and conceptual bounds of a paradigm of at least that component of the canon of scientific rationality composed of logico-empirical

---

19 Kuhn's conception of 'revolutions as changes of world view' is promoted ibid., pp. 111-35.
criteria of theory-evaluation. The independence of such a component from the conceptual context of any paradigm opens the possibility that there may exist rational grounds on which to choose to repudiate one's commitment to a paradigm. In the present model, such grounds are precisely of what the progressive faction in a revolutionary crisis avails itself in deciding to relax its commitment to the previously dominant paradigm: that faction believes, on paradigm-independent logico-empirical grounds, that the progress of science will be better served by the relaxation of the constraints on theory-choice imposed by the aesthetic canon. While Kuhn believes that the substitution of one paradigm by another in a scientific community is an entirely arational process, this model envisages that the abandonment of a paradigm may be decided by rational judgement. If aesthetic canons of theory-assessment constitute arational ruts which scientific progress marks out, requiring for a while that theories display certain aesthetic features which there are no rut-independent reasons to require, at least the desire to escape from the current rut admits sound rational justification.

5. Quarrels with the Kuhnian view of aesthetic factors

Although the present view of the dynamics of scientific revolutions has been expounded on the cues offered by the writings and terminology of Kuhn, to the extent that the previous sections were concerned largely with reconstructing Kuhnian concepts such as that of 'normal science' or of 'revolutionary crisis' in terms which appeal to the role therein of aesthetic theory-evaluation, there is sharp disagreement between this view of scientific revolution and that of Kuhn. Some areas of disagreement were pointed out during the above exposition. Another set of disagreements emerges however between this treatment and the brief remarks which Kuhn devotes to the operation of aesthetic features in science, and which have not yet been examined. These areas of contention are of interest as they illuminate both characteristic features of the present model of scientific rationality and the shortcomings which the present treatment perceives in other works. In this section, three lines of criticism of increasing generality will be moved against Kuhn's conception of the operation of aesthetic concerns in science.
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

The first concerns the role played by aesthetic considerations in times of paradigm-switch. Kuhn appears to believe that scientists will embrace a new paradigm at least partly in consequence of the appeal of the new paradigm’s aesthetic features, and that this transfer of allegiance will often occur despite the technical or empirical superiority of the pre-existing paradigm. Surveying possible ‘arguments for a new paradigm’, Kuhn writes that advocacy based upon the competing paradigms’ comparative ability to solve empirical or technical problems does not compel paradigm-switch, and that the reasons for which scientists adopt a new paradigm lie in a different class of arguments:

These are the arguments, rarely made entirely explicit, that appeal to the individual’s sense of the appropriate or the aesthetic – the new theory is said to be ‘neater,’ ‘more suitable,’ or ‘simpler’ than the old. 22

Having stated that arguments which draw their advocative force from the aesthetic features of a set of newly-formulated theories are apt to gain adherents to their paradigm, Kuhn suggests that such aesthetic appeal can play a role sufficiently important to determine the paradigm’s fate:

The importance of aesthetic considerations can sometimes be decisive. Though they often attract only a few scientists to a new theory, it is upon those few that its ultimate triumph may depend. If they had not quickly taken it up for highly individual reasons, the new candidate for paradigm might never have been sufficiently developed to attract the allegiance of the scientific community as a whole. 23

Kuhn stresses that on at least some occasions arguments based on aesthetic grounds counteract the weight of logico-empirical considerations which favour the continued adherence of the community to the pre-existing theories:

Something must make at least a few scientists feel that the new proposal is on the right track, and sometimes it is only personal and inarticulate aesthetic considerations that can do that. Men have been converted by them at times when most of the articulable technical arguments pointed the other way. When first introduced, neither Copernicus’ astronomical theory nor De Broglie’s theory of matter had many other significant grounds of appeal. 24

In the passages here reproduced Kuhn sketches the nature of the arguments which he believes weigh on one side and the other of debates on paradigm-choice. He expects that on at least some such occasions the mustered technical or

23 Ibid., p. 156.
24 Ibid., p. 158.
logico-empirical considerations will tell in favour of the retention of the pre-existing paradigm, whereas 'personal and inarticulate' aesthetic considerations will tend to prompt the adoption of the new paradigm. This view of Kuhn's is unequivocally contradicted by the proposal developed here, in which it is argued that in confrontations between an old and a new paradigm the logico-empirical arguments which bear upon the issue will tend to speak for the adoption of the new paradigm while the aesthetic considerations — possessing a conservative bias owed to the inductive manner of their construction — will recommend fidelity to the old paradigm. The balance of the historical evidence drawn from times of competition between long-established and novel paradigms strongly supports this treatment's attribution of aesthetic arguments to the conservative camp and of logico-empirical arguments to the progressive one in debates on theory-choice.

The difference of opinion with Kuhn does not render agreement with some of his historiographic judgements impossible, but it alters the significance to be read into some of the episodes on which he pronounces. For instance it is probably indubitable that, as Kuhn states in the last of the passages quoted above and as will be further investigated in chapter 10 below, Copernican theory found in aesthetic considerations a large part of its appeal to the community of mathematical astronomers. What is, on the contrary, subject to doubt is Kuhn's implicit interpretation of Copernican theory in that same passage as a revolutionary innovation, and consequently the legitimacy of his use of this episode to support his contention that revolutionary theories attract adherents largely through the appeal of their aesthetic features. The present treatment will maintain contrary to Kuhn that Copernican theory was not a revolutionary departure from the doctrinal context in which it was conceived, and further that it was able to attract adherents through the appeal of its aesthetic features precisely because of its aesthetic conservatism, its aesthetic conformity to the canons of pre-existing Ptolemaic astronomy.

Once the specific differences of opinion between Kuhn's view of the role of aesthetic considerations in theory-choice and that of the present treatment are set aside, a more deep-seated difference of approach to the notion of aesthetic factors in science emerges between the two treatments. Kuhn's brief writings on these factors are constrained by his rather inexplicably thinking of them as irremediably personal: he speaks of 'the individual's sense of [...] the aesthetic', of aesthetic considerations as 'highly individual reasons' for which to embrace a paradigm, of
aesthetic considerations as ‘subjective’, ‘personal’. It is not surprising that, as a methodologist keen to stress the social phenomena which determine and accompany paradigm-switch, Kuhn should devote less attention to any factor affecting theory-evaluation which could be shown to be truly individual, idiosyncratic to the point of evading social analysis. What is on the contrary surprising is that Kuhn should have come to regard aesthetic canons as the product of the taste of the individual scientist rather than as the construct of the community, which characterizes the community’s theory-preferences and determines the individual’s acts of theory-choice. The present treatment attributes the formulation of aesthetic canons not to the whim of the individual but to a communal metainduction over the performance of the community’s past theories, and suggests that this interpretation ensures that the phenomenon of aesthetic theory-assessment conforms to a Kuhn-like analysis of theory-choice to a far deeper degree than does the phenomenon of theory-assessment on logico-empirical criteria.

The third strand of criticism which the present treatment moves to Kuhn’s conception of the role of aesthetic criteria in science could equally be directed against virtually all models of scientific rationality proposed since the first appearance of Kuhn’s essay on scientific revolutions. None of these works appears to envisage the possibility that the construction or application of aesthetic criteria of theory-assessment may occur through a systematic mechanism, the study of which could hold lessons for the understanding of the dynamic of scientific progress as a whole. Where these works deign to allude to aesthetic evaluative criteria, they consign them to the category of unsystematic and inscrutable preferences. The abandonment of such an important component mechanism of scientific procedure to the margins of philosophical attention cannot but impoverish the accuracy or scope of models of scientific rationality, a limitation which the present treatment hopes to help overcome.

These expressions occur ibid., respectively on pp. 155, 156, 156 and 158. On the latter page Kuhn additionally calls aesthetic concerns ‘mystical’, a dismissal which cannot aid a rational or sensitive investigation of their genesis.
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

6. Boltzmann’s model of scientific revolutions

If the models of science propounded by philosophers have generally dedicated insufficient attention to the operation of aesthetic factors, one often encounters a greater awareness in those embraced by scientists. The expression of this awareness is frequently confined to unsystematic or fragmentary descriptions of the speaker’s own scientific practice, but in some more reflective writings scientists have delineated sophisticated views of the incidence upon scientific method of aesthetic considerations which extend to their role in scientific revolution.

One such model is contained in an article by L. Boltzmann which has been accorded insufficient attention by subsequent scholars. The paper comments incisively on several issues in scientific methodology, including the role of models in physics, the relative merits of realist and instrumentalist interpretations of theories, and the prospect of the underdetermination of theory-choice by data. It is however with Boltzmann’s view of scientific progress and of the nature of the revolutions by which that progress is punctuated that this section is concerned. That view accords closely with the treatment of scientific revolutions displayed in the present chapter, and it is of interest to this treatment to observe similar concerns aired by a practising scientist.

Boltzmann’s understanding of the history of science shows greater sophistication than was attained by the positivist views which prevailed in professional historiography for years even after his death. He first warns against too readily conceiving the mode of scientific progress as linear or continuous, and suggests that the past development of science does not appear to have been

---

26 The article is Boltzmann (1901). The sole discussion of it which I have found is Feuer (1974), pp. 335-41. It is neglected even by Scheibe (1988) – an otherwise praiseworthy treatment of the conceptions of scientific progress and revolution held by some nineteenth- and twentieth-century working physicists which discusses Boltzmann on pp. 141-5 – despite the fact that Boltzmann’s paper of 1901 plainly would bolster its argument. Scheibe’s understanding of Boltzmann’s views is apparently culled entirely from an obituary composed by him for J. Stefan in 1895.

27 Boltzmann’s remarks on the problem of underdetermination are brief but have a highly modern ring: ‘It is not inconceivable that two quite different theories should exist which are equally simple and which accord equally well with the phenomena, and which therefore, although they are totally different, are yet equally correct. The assertion that a given theory is the only correct one is merely the expression of our subjective conviction that there is no other theory so simple and according so well with the facts’ (1901, p. 244). Boltzmann appears to view this prospect as offering an argument for instrumentalism: ‘The object set us is not to discover an absolutely correct theory, but rather to light upon some constructive model which shall be as simple as the circumstances admit and represent the phenomena most adequately’ (ibid.).
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

guided exclusively by the application of logical criteria:

If we scrutinise the development of theory closely, it will be immediately apparent that its course has been by no means so continuous as we might have been inclined to believe, but rather that it is full of gaps and has not taken place, to appearances at least, along the simplest and most logical paths.\(^{28}\)

One expects any divergence of the history of science from 'the simplest and most logical paths' to be caused by scientists' adherence to extra-logical procedural criteria, by which scientists may have been tempted to wander from the theory-choices yielding the greatest empirical success. Boltzmann not only advances an explanation of this form for the divergence between the actual or historical development of science and its ideal or logical path, but also hints that the extra-empirical considerations which attract scientists may possess an aesthetic dimension:

Certain methods have frequently yielded the most beautiful results, and many persons have been tempted to believe that the development of science to the end of all time would consist in the systematic and unremitting application of them. But suddenly they begin to show indications of impotency, and all efforts are then bent upon discovering new and antagonistic methods. There then usually arises a conflict between the adherents of the old method and those of the new. The point of view of the former is characterised by its opponent as antiquated and obsolete; whilst its upholders in turn look down with scorn upon the innovators as perverters of true classical science.\(^{29}\)

Encapsulated in this brief passage is a view of revolution very similar to that of the present treatment, in which several stages of scientific procedure are discernible. Boltzmann ventures the suggestion that the apparent beauty of some theories would prompt scientists to advocate the continued application in future of the methods which yielded such theories, presumably in the hope that the same beauty would be reproduced in new theories. The present chapter suggested similarly that an aesthetic canon of theory-appraisal would entrench itself in the

\(^{28}\) Ibid., p. 229. Cf. Boltzmann's piece of 1895: 'The layman may have the idea that to the existing basic notions and basic causes of the phenomena gradually new notions and causes are added and that in this way our knowledge of nature undergoes a continuous development. This view, however, is erroneous, and the development of theoretical physics has always been one by leaps' (quoted in Scheibe 1988, p. 143).

\(^{29}\) Boltzmann (1901), ibid. Cf. again his treatment of 1895: 'In many cases it took decades or even more than a century to articulate fully a theory such that a clear picture of a certain class of phenomena was accomplished. But finally new phenomena became known which were incompatible with the theory; in vain was the attempt to assimilate the former to the latter. A struggle began between the adherents of the theory and the advocates of an entirely new conception until, eventually, the latter was generally accepted' (quoted in Scheibe 1988, ibid.).
procedures of a community, with the effect that scientists in the community would come to believe that all theory-choice should be decided in part by the application of the aesthetic canon to which they currently lend assent. Next, Boltzmann suggests, those previously apparently reliable methods show signs of empirical failure, a discovery which prompts a search for an alternative battery of methods or criteria. In the corresponding stage of the present model of scientific rationality, the crisis of theory-choice leads a faction of the community to relax its aesthetic commitments and embrace a new methodological canon shorn of aesthetic evaluative criteria. The progressive faction in this controversy regard the old theories explicitly as ‘obsolete’, the term used by Boltzmann to describe the corresponding judgement in his model; the adherents to the formerly dominant aesthetic canon see the new developments as a ‘perversion of true classical science’ or as a repudiation of cherished aesthetic commitments.

Boltzmann adds weight to the aesthetic dimension of the revolutionary rupture in science by likening it to similar discontinuities in fields of artistic endeavour:

This is a process, moreover, which is by no means restricted to theoretical physics, but to all appearances recurs in the history of every field of intellectual activity. [...] In like manner in art the Impressionists and Secessionists stand arrayed against the old schools of painting, and the Wagnerian school of music against the schools of the ancient classical masters. There is accordingly no occasion for surprise that theoretical physics does not form an exception to this general law.30

The interest of Boltzmann’s model of scientific development is heightened by the role which he attributes within it to himself. He writes not as a victor of history, a member perhaps of the progressive faction in a revolutionary crisis to whom history has accorded the privilege of scorning his opponents as ‘obsolete’. Boltzmann may well have witnessed a revolutionary crisis in theoretical physics, but he sees his part within it as having been that of the conservative partisan of the old doctrines repudiated by the progressive and eventually victorious faction:

When I look back over the manifold developments and transformation that have taken place, I seem to myself like a veteran on the field of science; nay, I might even say that I alone am left of those who embraced the old doctrines heart and soul; at least I am the only one who is still sturdily battling for them. [...] I appear before you, therefore, as a reactionary and belated thinker, as a zealous champion of the old classical doctrines as opposed to the new.31

30 Boltzmann (1901), pp. 229-30.
31 Ibid., p. 233.
CHAPTER EIGHT: REVOLUTION AS AESTHETIC RUPTURE

If Whig history is ‘the tendency [...] to praise revolutions provided they have been successful, to emphasize certain principles of progress in the past and to produce a story which is the ratification if not the glorification of the present’, Boltzmann’s melancholy testament offers a valuable insight into the concerns of the defeated. It is a matter for admiration that the disillusionment which Boltzmann must have suffered in his work in physics did not prevent him late in life from holding a view of scientific revolutions which captures the most progressive and innovatory impulses in scientific methodology.

7. The analogy with revolutions in art

Boltzmann suggested in one of the passages cited in the previous section that the development of the sciences and the arts is punctuated by revolutionary discontinuities of the same form. Analogies have frequently been drawn between the manner of the progress traced by artistic and scientific disciplines. The problem which has bedevilled most such analogies has been their lack of specificity, which has prevented determinate and illuminating conclusions from being drawn about the dynamic of either of the *relata*. The lack of specificity has in turn been occasioned by the unavailability of determinate models of revolution in either the sciences or the arts, able to predict the sequence of phases within revolution more determinate than stages labelled generically ‘loss of consensus’ or ‘break with the past’.

The construction earlier in this chapter of a highly determinate model of revolutions in science restores interest to an intellectual historian’s revisiting the issue. This section will briefly sketch an analogy between revolutions in science as interpreted above and revolutions in certain art forms.

The model of scientific progress developed in the present treatment pays regard to the development of three entities within science: the empirical power of the sequence of theories embraced by a community, the aesthetic features displayed by those theories, and the aesthetic evaluative canon of the community. Briefly, continuous development will be obtained while the evolution of the aesthetic canon of the community roughly maintains pace with the evolution of

---

the aesthetic features exhibited by the sequence of theories embraced by the community; if, however, theory-succession governed by empirical criteria communicates too high a rate of evolution to the aesthetic features of the community's set of theories, the community's aesthetic evaluative canon will fail to accommodate those changes: in order to restore accord in theory-choice the community will be forced to exchange its old aesthetic evaluative canon for a new one.

Analogies are readily suggested between the operation of this mechanism in science and the development of art forms characterized by both aesthetic canons and technical capability. Examples of such art forms are architecture and industrial design.\(^{33}\)

In architecture a certain material or technique of construction fosters the establishment of a certain aesthetic canon. For instance, the technological exploitation of steel and plate glass is clearly both necessary to and supportive of the aesthetic canon of modernism: it is necessary for that canon since modernist projects could not have been realized without the use of steel and plate glass, and it is supportive of that canon since given such technological means the community of architects will naturally be driven to exploit them in buildings which utilize them in aesthetically distinctive ways. In time the aesthetic canon thus constructed will assume normative functions, and buildings which fail to satisfy the canon will be criticized on those grounds. Aesthetic canons thus emerge in architecture at least partly from technological capability. One may speak literally of the aesthetic of, say, reinforced concrete, meaning by that expression the aesthetic canon which naturally springs from the architectural use of that material.\(^{34}\) The aesthetic which arises from the use of a certain constructional technique is analogous to the aesthetic canons which arise from the empirical success of scientific theories exhibiting certain aesthetic features. In each case the community's technical capability supports or fosters the formulation of specific aesthetic canons and their subsequent prescriptive imposition upon the

---

\(^{33}\) Strictly speaking the development of any art form, not only that of forms like architecture or industrial design, is determined partly by advances in relevant technical capability: the aesthetic canons of oil painting or of the performance of music are affected by innovations in the technology of pigments and musical instruments. For instance the evolution of the aesthetic canons of keyboard music from Bach to Beethoven has been determined partly by innovations in the harpsichord and piano. However the effects of advances in technical capabilities is more readily apparent in art forms which make more crucial use of technology: the succeeding treatment will refer to architecture as one such form.

\(^{34}\) An example of a textbook which lays out the architectural aesthetic canon associated with a particular material or technique is Michaelis (1963).
practice which first achieved that technical capability.

Naturally architecture witnesses a constant evolution of its technical capabilities: new materials and techniques are constantly devised. One imagines that if this rate of technical innovation is sufficiently low, it will be matched by the rate of evolution of the community’s aesthetic canon: similarly in science if the rate of aesthetic innovativeness of the sequence of theories embraced by the community is low, the community’s aesthetic canon will maintain its accord with current theory-choice. If however a new material or technique is suddenly introduced which affords architects radically different aesthetic possibilities, the community may well perform a discontinuous exchange of its aesthetic canon for a new one as it turns to explore the constructional possibilities opened by the new technique. This procedure is closely analogous to the scientific community’s exchange of one aesthetic canon by another which better accords with current, empirically-driven theory-choice.

8. The analogy with political revolutions

Kuhn’s imagery strongly evoked analogies between revolution in science and in society.\(^\text{35}\) This correlation was not altogether original to his work: the connections between discontinuities in science and in social formations have been traced since the mid-seventeenth century.\(^\text{36}\) As in the case of connections drawn between scientific and artistic revolutions, the chief difficulty attendant upon such analogies is the achievement of a degree of determinacy sufficiently great to ensure interest. The problem is alleviated in this case by the existence of a determinate model of political revolutions, that proposed by Marxian political economics. That model is a promising candidate as an analogical *relatum* of the present model of scientific revolutions, for it correlates the development of the technical capabilities of a community on the one hand and the evolution of constructs which derive from those capabilities on the other.

Marx’s theory of economic history locates the technical capabilities of a

---

\(^{35}\) Kuhn (1962), e.g. pp. 92-4.

\(^{36}\) For notes on the history of the analogy, see I.B. Cohen (1985), pp. 7-14, 473-7 and *passim*; Feuer (1974, pp. 252-68) interestingly explores the disanalogies of political and scientific revolutions.
community in an economic substructure which leads to the establishment of a particular social formation. Generally in history, the social formation given to itslf by a community is appropriate to the current state of the community's technical capabilities, in that it permits the full and efficient exploitation of those capabilities to the benefit of members of the community. But developments in the available technical means continue apace. The social formation typically possesses sufficient flexibility that its structure and institutions can accommodate the exploitation of some new technical capabilities; in time however the members of the community will discover that full exploitation of newly-developed technical capabilities is impossible within the existing social structure. The social formation has now begun to hamper the technical progress of the community. In the resulting crisis the old and counter-productive social structure is overthrown and replaced by one more attuned to the new technical means. Thus the economic substructure engenders, consolidates, is gradually hampered by and ultimately overthrows and replaces a superstructure composed of, in Marx's term, relations of production.\textsuperscript{37}

One may in science analogously conceive of empirical progress constituted by increasing empirical power as the substructural motor which accommodates successive superstructures consisting of aesthetic canons. As empirical progress is continuously achieved, the prevailing aesthetic canons will be first strengthened or entrenched by the accord which is manifested between them and the empirical evaluative canon. If however the empirical progress of the community's theories is too swift, tensions will arise between the theory-choices of the aesthetic and the logico-empirical canons which will undermine and ultimately destroy the prevailing aesthetic canon. The periodic overthrow of successive aesthetic canons is due to the fact that soon after their formulation they switch to a conservative evaluative tack and in time inevitably begin to hamper scientific progress, exactly as state institutions once erected crystallize a fleeting order and thus tend to hinder continued social evolution.

\textsuperscript{37} A detailed description of this model of economic history and revolution is contained in G.A. Cohen (1978).
Chapter Nine
IMPLICATIONS FOR HISTORIOGRAPHY

1. The possibility of writing history

The present model of scientific revolutions, which identifies such events with discontinuous changes in communities’ aesthetic canons of theory-assessment, holds some important consequences for the practice of the historiography of science. In this chapter, two historiographic consequences will be discussed. The next section will analyze what suggestions the model advances on the method by which the occurrence of a revolution may be discerned in the historical record; the current section will address the prior question of the extent to which it is at all possible to write the history of a science of which the progression is punctuated by revolutions. The discussion will include comparisons between the stance taken on each topic by Kuhn and that of the present treatment.

Kuhn’s views on the occurrence of scientific revolutions stand in an ambivalent relationship to historiographic investigation: if, on the one hand, he proclaims that studies of science would be transformed by a more accurate portrayal of the history of its practice, on the other the characteristics which he attributes to scientific revolutions appear to preclude the possibility of answering many of the most interesting classes of questions which the discipline of history of science has taken it upon itself of posing both before and since the publication of Kuhn’s monograph. Thus a thoroughgoing acceptance of Kuhn’s views would tend to impede the accomplishment of the project which those views prescribe. It may be for this reason that Kuhn’s two extended works of historiography, which seek to reconstruct episodes of theory-succession which were in his opinion revolutionary, appeal to a notion of ‘scientific revolution’ no more laden with the theoretical apparatus of his model of revolutions than is the notion of revolution invoked by many a philosophically untutored treatise in history of science.

The difficulty which in Kuhn’s model would attach itself to the work of the historian of science derives in part from his insistence that successive paradigms
are separated by instances of radical standard variance. Consequences of this phenomenon for analytical philosophy of science were reviewed in the previous chapter, but its historiographic implications were there devoted scarce attention.

Because of standard variance, those relationships between reasons for and outcomes of theory-choice which hold in the canon of rationality characteristic of the scientist of one paradigm may fail to hold in the canon proper to the historian in a later paradigm: in particular, a reason which in the former paradigm will suffice to determine the choice of a certain theory over its competitors may prove inadequate to force the similar choice in the later paradigm. In consequence of the gradual opening of this inferential gap, even if the arguments of scientists constitute a sufficient explanation of their theory-choices within their own paradigm, they may not hold in the paradigm successive to their own. What is a sufficient argument to compel theory-choice in its own paradigm becomes an argument insufficient to communicate an understanding of that theory-choice in a paradigm which follows. Hence on Kuhn’s model even (or especially) a historian versed in the theories of the modern paradigm will be unable satisfactorily to understand episodes of theory-choice in a paradigm of the past.

The conclusion appears to preclude the success of the historian’s work or, at least, to circumscribe its scope. This observation counts in some respects to the disfavour of Kuhn’s model: conceptual models generally gain favour if they extend the scope of philosophical enquiry or demonstrate the possibility of its success. The model of scientific methodology contained in the present treatment does not purport to show that the past practice of science is opaque to historiographic understanding to the extent to which Kuhn believes it to be. The greater scope which the present treatment attributes to historiography is opened by its assertion that only some of the criteria of theory-choice of a community vary at the exchange of one paradigm for another, while another component of the overall canon of criteria persists unchanged across paradigm boundaries. More specifically, the present treatment believes that at times of revolution only the canon of aesthetic criteria of theory-evaluation changes while the battery of logico-empirical criteria remains unaltered.

As a consequence of the partial invariance of criteria of theory-assessment, all paradigms share one of the two components of science’s overall canon of theory-appraisal. What constitutes a strong logico-empirical reason for preferring one theory to another in a given paradigm will retain its strength in the paradigms of successive times. Hence at least part of scientists’ overall grounds for preferring one theory to another will be as readily comprehensible and as apt.
CHAPTER NINE: IMPLICATIONS FOR HISTORIOGRAPHY

to be judged conclusive to later historians as they seem to the very scientists who first made that choice. It is for this reason that the splitting of criteria of theory-evaluation into two classes exhibiting different historical behaviour renders the past record of science more accessible to historiographic understanding than Kuhn maintained.

For sure, the identification of revolutions with discontinuous changes in aesthetic canons of theory-assessment vindicates Kuhn’s warning that the reconstruction of instances of theory-choice in paradigms previous to our own may be arduous. After all, we no longer acknowledge the aesthetic canons which defined those paradigms and contributed to decide the theory-choices performed within them: we have constructed in their stead our own aesthetic canons by metainduction over the performance of theories developed between the time of the paradigm under reconstruction and today. For example, we cannot credit as justified Copernicus’s predilection for the circle as the geometrical form of planetary orbits because there is not now a commitment to Pythagorean simplicity or to Aristotelian natural philosophy in our aesthetic canon for the appraisal of theories in mathematical science; again, we dismiss as unwarranted much of Einstein’s opposition to quantum mechanics because we no longer count among our metaphysical commitments the tenet that physical theories should be deterministic. Nonetheless, the reconstruction of past instances of theory-preference is not thereby rendered impossible, for we share the logico-empirical criteria of theory-evaluation proper to past paradigms.

The chief task for the historian of scientific revolutions is therefore that of identifying the aesthetic canons of a paradigm in pre-revolutionary times and of charting the theory-choices performed in accordance with those canons which depressed the observational success of the theories of that paradigm and ultimately led to its demise.

2. The recognition of scientific revolutions

The model of scientific revolutions presented in the previous chapter implicitly contains precise advice for the recognition of the occurrence of a revolution in the historical record. If a historical interval has witnessed the adoption of a succession of theories by a community, and the historian determines the nature of the
CHAPTER NINE: IMPLICATIONS FOR HISTORIOGRAPHY

grounds considered by the community to have weighed for and against the adoption of each of those theories, then the occurrence of a scientific revolution within that interval is revealed in a temporary change (to be described below) in the nature of those grounds, and the theory at the time of which that change took place is the revolutionary innovation. The change in the grounds for theory-adoption which is associated with a revolution occurs as follows.

According to the previous chapter, a non-revolutionary theoretical innovation will be able to gather adherents in the community both on the strength of its empirical success and in virtue of its aesthetic features. This is because an innovation which falls short of effecting a revolutionary rupture with the past will continue to adhere to the pre-existing aesthetic canons of the community: the community will therefore value a theory which constitutes a non-revolutionary innovation by dint of its continued adherence to the extant aesthetic canons as well as on the strength of such empirical success as it may achieve. On the contrary, a theoretical innovation which effects a revolutionary innovation may certainly gather adherents by reason of its empirical success, but will generally be unable to attract endorsement in virtue of its aesthetic features. This is because a revolutionary theory – on account of the above definition of what constitutes a revolutionary innovation – will violate the existing aesthetic canons of the community and thus be perceived at least at first as aesthetically displeasing.

As a consequence, in order to decide whether on the definition developed here a certain theory constituted at the time of its proposal a revolutionary innovation or an innovation which effected a shallower rupture with past science, it is helpful to examine the nature of the grounds considered by members of the community of the time to weigh in favour of and against the adoption of that theory. If the historian perceives that a community comes to adhere to a new theory entirely or partly on aesthetic grounds – i.e. on the strength of that theory's accord with the community’s pre-existing aesthetic canon of theory-evaluation – then the introduction of that theory will not have constituted on the present analysis a revolutionary departure from pre-existing theoretical tenets. On the other hand, if members of the community adhere to the new theory on the grounds of its perceived empirical success but express displeasure or unease at its aesthetic features, or state that they find it necessary to suspend their faculties of aesthetic theory-evaluation in order that they may be allowed without inner conflict to reap the predictive benefits of the adoption of the new theory, or if the same theory is embraced by one faction of the community on empirical grounds
but resisted by another on aesthetic grounds, then the present analysis interprets that theory to have effected a revolutionary break with the pre-existing aesthetic canons of the community and thereby with earlier scientific practice as a whole. By examining which of these two cases holds for a certain past theory, the historian may determine whether on the present account that theory initiated or not a scientific revolution. The signs of a scientific revolution in a community are thus visible in the nature of the community’s grounds for theory-adoption.

Thus, on the present model of scientific revolutions, the occurrence of a revolution can be discerned in the historical record by examining only historical evidence concerning the practice of science at the time of the revolution itself, and in particular evidence about the grounds considered by the community at the time to have weighed for or against the adoption of the theories then being proposed. On the present treatment, historiographic reports about the nature of the grounds on which a past theory gathered adherents in its community indicate whether that theory constituted a revolutionary or non-revolutionary innovation in its community. In other terms, the task of detecting the occurrence of a revolution in the historical record is on this model accomplished by reference to data drawn solely from within the historical record itself. On the present model this task requires no non-historiographic consideration of the theories concerned: it requires, for instance, no retrospective investigation of the logical interdependence of successive pairs of the theories proposed during the historical interval being examined for the occurrence of a revolution.

That on this model the historian can – purely as historian – detect the occurrence of a revolution in the historical record is a distinctive feature of this model, since it is shared neither necessarily nor – as will shortly be seen – in fact by other models of scientific revolutions. It is furthermore a feature of some interest, since it renders the project of discerning the occurrence of a past revolution less subject to retrospective or ahistoricist judgement: if the project is to be accomplished by reference to data from the historical record alone, rather than by appeal to subsequent rationalizations of scientific method, the resulting judgements are likely to achieve greater historiographic authenticity.

The value of being able to detect the occurrence of a scientific revolution by reference to data from the historical record alone is best revealed by considering the manner in which the same task is accomplished on the Kuhnian model of revolutions. Kuhn believes that ‘the transition from Newtonian to
Einsteinian mechanics illustrates with particular clarity the scientific revolution.\(^1\) What according to Kuhn is the nature of the evidence for supposing that these two contributions to physical science are indeed separated in the historical record by a revolution? Kuhn believes that the required evidence is to be found in ‘the relation between contemporary Einsteinian dynamics and the older dynamical equations that descend from Newton’s *Principia*.\(^2\) The particular relation that Kuhn believes is indicative of the occurrence of a revolution appears to be one of logical inconsistency:

> From the viewpoint of this essay these two theories are fundamentally incompatible in the sense illustrated by the relation of Copernican to Ptolemaic astronomy: Einstein’s theory can be accepted only with the recognition that Newton’s was wrong.\(^3\)

Kuhn drives home this point by tackling the positivist counterargument to his view that Newtonian and Einsteinian dynamics cannot have been separated by a revolutionary discontinuity since Newtonian dynamics is a special case of Einsteinian dynamics and can be inferred from the later theory for the particular conditions in which the velocities are much smaller than that of light. Kuhn rebuts to this objection that the physical referents of the Einsteinian notions of space, time and mass are not identical from those of the Newtonian notions which bear the same name; hence ‘we cannot properly be said to have derived Newton’s Laws.’\(^4\) For Kuhn the fact that the derivation fails to carry is a demonstration of the logical incompatibility of the two theories, and this in turn is as conclusive an indication as may be required for the proposition that a revolution separated the Einsteinian theory of dynamics from that of Newton.

For all Kuhn’s valuable insistence on the value of historiographic insight in the philosophy of science, it is clear that the nature of the evidence which Kuhn takes to support the suggestion that Einsteinian physics is separated from Newtonian physics by a revolution is profoundly ahistoricist. The relation of non-deducibility between the two theories which he takes as indicative was not a concern of the community of physical scientists at the time of Einstein: if that community perceived itself to be traversing a revolution, this perception was not held in virtue of the observation that from Einstein’s theory it was impossible to

---

\(^1\) Kuhn (1962), p. 102.
\(^2\) Ibid., p. 98.
\(^3\) Ibid.
\(^4\) Ibid., p. 102; emphasis in the original.

157
derive Newton's theory as a special case.\textsuperscript{5} Thus, in order to discern the occurrence of a revolution in the historical record Kuhn relies on evidence which, far from being historiographic in origin, pertains to observations on intertheory relations which were not pursued in the communities caught in the revolutions which Kuhn claims thereby to have detected. By contrast, the present treatment proposes to discern the occurrence of a revolution on the basis of evidence gleaned from the historical record itself: this evidence concerns judgements passed by the communities themselves, and concerns in particular the grounds on which the communities justified their adoption or rejection of theories.

\textsuperscript{5} On the contrary, physicists from the time of Einstein to the present day have seen it as a reason for satisfaction in the state of the science of mechanics that Newton's laws emerge from relativistic mechanics as an approximation under special conditions.
Chapter Ten
APPLICATIONS OF THE MODEL TO HISTORY OF SCIENCE

Kirk was as much overcome by the beauty, simplicity and economy of this solution as Copernicus must have been when he first thought of putting the sun in the centre of the Solar System and saw all the planets, instead of describing complicated and ugly geometrical capers, move onward in orderly and dignified circles. He sat and contemplated it with affection for nearly ten minutes before venturing to examine it. He was afraid of knocking the bloom off it.¹

Dorothy L. Sayers

1. The test of the model

The conceptual field is now set for the performance of a war game. The beginning of this treatment likened part of the work of the philosopher of science to the re-staging of episodes of military history in which the actions of characters personating the original protagonists are directed by a tactical canon compiled by the organizers of the war game. Case studies in scientific methodology are similarly the arena in which the methodologist's proposed inferential canon directs the actions of characters representing or personating historical scientists. Such case studies aim to re-enact episodes of scientific history in which it is presumed that the proposed inferential canon played a determining part, and thereby to test the inferential canon as a model of the original participants' canon of rationality by examining the agreement of the re-enactment with historical data independently gathered on the episode.

The past chapters have sought to elaborate an inferential canon for the operations of theory-assessment and choice: the canon developed there appeals distinctively to aesthetic evaluative criteria. The treatment has already advanced on behalf of this canon both systematic justification, drawn from aesthetic theory and considerations of scientific methodology, and historiographic justification, aiming to demonstrate scientists' use of the canon in their work. A systematic

case study – the metaphorical equivalent of a pitched war game – has however lacked.

Some brief considerations are able to guide the choice among historical episodes to those most suited to reveal the operation of aesthetic canons of theory-assessment. The previous chapter has made it clear that the juncture at which aesthetic criteria of theory-evaluation assume the most critical role – and therefore also the phase of scientific progress in which most promisingly to study their operation – is the scientific revolution. The reason is the following. A non-revolutionary period in a scientific discipline is defined – on the account of chapter 8 – as a time-span in which the empirically successful theories of that discipline conform to a common aesthetic canon, or in other terms preserve aesthetic continuity with each other and with their immediate predecessors in the period. Such a period is obtained in the history of a discipline for as long as, as a matter of contingent fact, the aesthetic metainduction yields an aesthetic evaluative canon which attributes favour to the aesthetic features embodied by the empirically most successful of the theories newly formulated in that discipline, i.e. for as long as the aesthetic evaluative canon of the community continues to recommend in acts of theory-choice the adoption of the same theories to which preference is accorded by the community’s logico-empirical evaluative criteria. In that time-span theory-evaluation is thus non-controversial, as the logico-empirical and the aesthetic evaluative canons will agree in their recommendations for theory-choice.

A scientific revolution occurs on the contrary when theory-choices performed on logico-empirical criteria depart from theory-choices performed on aesthetic criteria, following the revolutionary emergence of new theories which do not conform to the hitherto dominant aesthetic canon but demonstrate empirical success greater than the success of those of their competitors which do so conform. Theory-choice will then become controversial, since the theory-evaluations performed on empirical criteria will disagree with those performed under the dominant aesthetic canon. It is at such times that aesthetic criteria – so to speak – escape from the shadow of empirical criteria, and yield different recommendations for theory-choice. The junctures at which one may most sensitively discern the operation of aesthetic criteria of theory-evaluation are thus times of revolution, at which one will expect them to play a conservative role in advocating the retention of theories conforming to the hitherto dominant aesthetic or stylistic canon and the rejection of their new, aesthetically-innovative competitors.
As an application of these historiographic principles, and in order to test the capacity of the present model to account for past scientific practice, the remainder of this chapter will examine two pairs of episodes in the history of science. The first pair - of which the discussion will extend over sections 2 and 3 - is constituted by the Copernican and Keplerian contributions to planetary astronomy, the second - of which a survey will occupy sections 4 and 5 - by Einstein's enunciation of special relativity and the development of quantum mechanics. That these episodes are extensively discussed in extant historiographic literature adds to the interest, for the interpretations to follow will cast doubt on some such treatments.

The first episode of each pair is commonly considered revolutionary: one speaks naturally of a 'Copernican revolution' and of the revolutionary impact of special relativity. Under the perspective erected in the previous chapters, however, both these innovations preserved continuity with decisive elements of pre-existing aesthetic canons: briefly put, Copernicanism fulfilled the existing metaphysically-justified requirements of simplicity (principally the notion that the circle was the elemental path of celestial bodies, not necessitating explanation), while special relativity satisfied the requirements - generated by nineteenth-century physics - of theoretical symmetry and of consistency with metaphysical commitments (viz., the tenet of the compatibility of mechanics and electromagnetism and the doctrine of causal determinism). It will be argued below that because these innovations retained such significant continuity with pre-existing aesthetic canons, neither constituted a scientific revolution in the sense of that notion elaborated in the present treatment, which interprets a revolution as the abandonment of commitments to a pre-existing aesthetic canon. It will be argued that on the contrary Copernicanism is most appropriately seen as the culmination of the programme of Ptolemaic astronomy (which, in conjunction with Aristotelian natural philosophy, enunciated and entrenched the commitment to the notion that circular motions were the sole proper for celestial bodies) and special relativity as the culmination of the programme of nineteenth-century physics.

Each of these two episodes will be contrasted with the second member of the pair of episodes to which it belongs, which constituted a genuine rupture of aesthetic canon and hence, on the present view, a revolution: the Ptolemaic commitment to circular celestial motions was broken by Keplerian introduction of ellipses and the Einsteinian commitment to determinism was relaxed by those who developed quantum mechanics. There will further be adduced evidence that
the scientists who in each episode retained a commitment to the pre-existing paradigm explicitly regarded the revolutionary theories of their time as aesthetically displeasing: Ptolemaic or strict Copernican astronomers abhorred Kepler’s ellipses as imperfect and improper to the celestial domain just as Planck and Einstein considered quantum mechanics aesthetically repugnant in consequence of its indeterminism.

The treatment to follow will therefore address the problem of discerning rupture in the historical record: it will deny that certain innovations constituted a rupture in scientific progress and assert that on the contrary certain other developments represented breaks of the stipulated kind. The task might reduce to the banal enterprise of tidying and labelling facts on the chronological axis were the project not guided by a theory of the dynamics of scientific revolutions.2

2. Aspects of continuity: Copernicus’s contribution to astronomy

That there occurred a revolution in mathematical astronomy at some time during the sixteenth or seventeenth centuries appears highly plausible: if the notion of a discontinuity in theory-succession has any application, one such fracture surely intervened between the Ptolemaic conception which dominated Western astronomy in 1500 and the theoretical scene upon which strode Newton.3 The issue for philosophical historiography is to discern which of the theoretical innovations which took place during this span were of sufficient moment and of the appropriate nature to have constituted a revolutionary discontinuity.

Since the mid-eighteenth century it has been customary to settle this issue by associating a revolution with the work of Copernicus.4 Even Kuhn, whilst

---

2 The four studies to follow do not presume to bring to light any previously unknown historical fact, but rather – true to their nature of rational reconstructions – seek to ascertain the degree of coherence between a model of scientific practice and what is known independently of that model about certain historical episodes. Furthermore, it would clearly have been impossible to refer to even a substantial proportion of the available secondary literature on such episodes as the reception of Keplerian astronomy or of the theory of special relativity: judgements of the relative importance and relevance of material play a large part in determining the content of the remainder of this chapter.

3 For advice on the topic of this section and the next I am grateful to Dr J.L. Russell of Heythrop College, University of London.

somewhat qualifying the degree of originality which he attributes to Copernican astronomy by raising the question 'whether Copernicus is really the last of the ancient or the first of the modern astronomers', yet nonetheless accepts the customary view: 'The publication of Copernicus' De Revolutionibus Orbium Caelestium in 1543 inaugurates the upheaval in astronomical thought that we call the Copernican Revolution.'

A historian approaching such long-standing unanimity runs the risk of supposing that the characterization of Copernicus's work as revolutionary is a brute or incontrovertible historiographic fact, and of forgetting that the register of events which are taken to count as scientific revolutions is determined by the philosophical theory against which one views the historical record. The present treatment approaches historiography armed with a theory of revolutions different from that which bequeathed the notion of 'the Copernican revolution': it must hence settle on its own judgement of the depth of the innovation constituted by the work of Copernicus.

As was intimated in the previous section, this judgement will be formed by investigating the nature of the grounds on which the Copernican theory was embraced. Before however that investigation in theory-adoption is undertaken, it is necessary to limit its scope by drawing a disciplinary boundary, distinguishing the innovative effect of Copernicanism in mathematical astronomy from its effect in general learning. Much conventional historiography saw Copernican astronomy as revolutionary in virtue of its most eye-catching innovation, the displacement of the earth from the centre of the universe; and there is no doubt that this change exercised a revolutionary effect in wide intellectual matters both at the time of its first broaching and for centuries to follow. But the history of mathematical astronomy is concerned not primarily with such evolution in general mentality but rather with innovations in the technical elements of theories, those aiming at the prediction of observational data or the elucidation of the physical configuration of the universe. The mere switch from geocentrism to heliocentrism

---

5 Kuhn (1957), p. 182.

6 Ibid., p. 134. Kuhn transcribes the last word of the short title as Caelestium instead of the more usual form, Coelestium. Kuhn emphasizes the purportedly revolutionary character of Copernicus's contribution still further in his later monograph (1962, e.g. pp. 149-50), which, as perhaps befits a work of philosophy of science rather than of historiography, treats the phenomenon of scientific revolution at a higher level of generality.

7 Two hundred years later for instance Kant took the 'Copernican revolution' as the archetypal example of an intellectual development which had the effect of reappraising the relation between the human mental or cognitive faculties and external reality.
can be counted as one of the innovations wrought by Copernicus in general learning, but not in technical planetary astronomy. What differences there were between the predictive consequences of Ptolemaic theory and those of Copernican theory were not due simply to the difference between the identity of the bodies which occupy the centres of the two systems: they are due instead to the difference between the specific arrangements of circles freely chosen in each of the two theories within the constraints of geocentrism and heliocentrism respectively. The mere substitution of heliocentrism for geocentrism cannot hence of itself count as an innovation introduced by Copernicanism within the domain of mathematical planetary astronomy.

The distinction between its effect in mathematical science and its repercussions in general mentality locates much of what is commonly considered to be the revolutionary impact of Copernicanism outside the boundaries of planetary astronomy: this distinction thus greatly reduces the degree to which within the discipline the work of Copernicus can be considered revolutionary. The determination of this degree will be completed now by the method intimated in section 1 of this chapter, i.e. by an analysis of the nature of the grounds considered by the early adherents to the Copernican theory to have weighed in favour or against its adoption.

Classical historians have typically assumed that Copernicus's theory gained adherents on the strength of its predictive accuracy and its degree of simplicity, and in particular that it won adherents from the pre-existing Ptolemaic theory in virtue of its possession of these two features to a degree greater than that of its

---

8 Evidence for the proposition that a switch from geocentrism to heliocentrism is not sufficient to determine changes in the predictive consequences of planetary theory is contained in the following observation: that it is not the case that any geocentric theory will yield predictions different from those yielded by any heliocentric theory. This in turn is demonstrated by a comparison between the heliocentric Copernican and the geocentric Tychonic systems. These are mathematically equivalent and so yield precisely the same empirical predictions. Hence there is at least one geocentric theory which yields the same predictions of planetary positions as a heliocentric theory. On the Tychonic system and its mathematical equivalence to the Copernican system see e.g. Kuhn (1957), pp. 201-4. The switch from geocentrism to heliocentrism effected by Copernicus differs in this respect from the transition from circular orbits to non-circular, elliptical orbits promulgated by Kepler: the latter change, unlike the former, is a condition sufficient to determine changes in the predictive consequences of planetary theory.

9 The distinction between the domains of philosophical cosmology and of mathematical astronomy is further discussed in Hanson (1961), pp. 169-72.
Chapter 8 showed how a theory which constitutes a revolutionary innovation in its discipline will be unable to gather support on the strength of its aesthetic features since these will be perceived to violate the then-dominant aesthetic canons of theory-appraisal. Thus the sole features in virtue of the possession of which a theory that constitutes a revolutionary innovation will be able to gather endorsement will be logico-empirical. Now, the (high or relatively high) empirical adequacy and degree of simplicity to which classical historiography attributes the Copernican theory’s attraction of adherents are on the present account just such logico-empirical features. Thus, the report by classical historiography that Copernican theory gathered supporters in virtue of its possession of such features is consistent with, and may suggest the truth of, the proposition that Copernican theory constituted a revolutionary innovation in its discipline. But could Copernican theory possibly have gathered support on the strength of the degree to which it possessed empirical adequacy and simplicity?

First, predictive accuracy. There are two classes of predictions for which the Copernican theory might hope to claim an accuracy superior to that of the Ptolemaic: quantitative predictions of the positions of the planets, and qualitative predictions of the appearance of the night sky to the earth-bound observer.

Copernicus pronounces himself content in the *Commentariolus* with the accuracy of the predictions of planetary positions yielded by Ptolemaic theory, so it does not appear at least that he was impelled in his research by an urge to improve on the accuracy of the quantitative predictions of astronomical theory. The quantitative predictions of the Ptolemaic and the Copernican theories have been compared by several historians of mathematical astronomy, and – as if to confirm the expectation raised by the previous sentence – the latter have been found no more accurate than the former. Furthermore, to decide between the Ptolemaic and Copernican models on the basis of accuracy of prediction would have required more precise data than existed at the time of Copernicus, or for

---

10 Examples of histories of astronomy which make this assumption are given in the text and notes below. In a somewhat more exhaustive treatment of this issue than it is possible to give here, Palter (1970, p. 105) discusses the comparative merits of the Ptolemaic and Copernican theories in terms of three criteria: simplicity, accuracy, and physical plausibility. He argues that to gauge the appeal of Copernicanism in such terms is historiographically legitimate, in that these criteria for the assessment of astronomical theory were not introduced retrospectively by historians but prescribed by practising astronomers from the time of Eudoxus to that of Copernicus.

11 Copernicus states in the *Commentariolus* that ‘the theories concerning these matters that have been put forth far and wide by Ptolemy and most others [...] correspond numerically [with the apparent motions]’ (Swerdlow ed. 1973, p. 434).

decades afterwards. Thus even if the Copernican model had yielded quantitative predictions more accurate than its Ptolemaic competitor, this superiority would not have been apparent to the discussants at the time and hence could not have weighed as a factor determining theory-choice.13

The alleged superior accuracy of the Copernican theory manifests itself no more strongly in the domain of qualitative predictions about the appearance of the solar system. A comparative test of the accuracy of the qualitative predictions of the Copernican and Ptolemaic theories is offered by the observed approximate constancy of the brightness of Venus during the course of its and the earth's orbits. As both Copernicus and Osiander note in the De revolutionibus, the difference between the maximum and the minimum values attributed by Ptolemaic theory to the distance of Venus from the earth is comparatively very large. One would expect that if the universe were as Ptolemaic theory said, the brightness of Venus as observed from the earth would undergo large variations. While the variation in the apparent brightness of the other planets accords reasonably closely with the variations attributed by Ptolemaic theory to the distances between those planets and the earth, there is no such close accord in the case of Venus: the apparent brightness of Venus varies surprisingly little. This observation counts in principle as a failure of one of Ptolemaic theory's qualitative predictions of the appearance of the night sky, and hence — in a case of comparative theory-evaluation such as this is — as an instance supporting Copernican theory. Copernicus and Osiander hail the observational report presumably as such a confirming instance. But the fact which they seem to neglect is that Ptolemaic theory and Copernican theory happen to predict the same values for the fluctuations of the distance between Venus and the earth, and that the Copernican theory offers no separate explanation for the observed constancy of the Venusian luminosity. So the extent to which the observational report disconfirmed the Ptolemaic theory was roughly the extent to which it disconfirmed the Copernican theory.14

13 The extreme historiographic thesis that by the mid-sixteenth century the Ptolemaic theory was in a state of generalized 'empirical crisis', which Copernicus resolved, is examined and rejected by e.g. Burtt (1955, pp. 36-56) and Gingerich (1975).

14 For further details see Price (1959), pp. 212-4. The true explanation for the constancy of the apparent luminosity of Venus lies in a coincidental compensation between the effects on its brightness due to its distance from the earth and those due to its phases; this explanation was attained only with Galileo's telescope observations. Galileo discusses the matter in his (1632), pp. 328, 339, where he expresses astonishment at Copernicus's seeming failure to appreciate the fact that his own theory — as well as Ptolemy's — entailed the prediction of a wide variation in the apparent luminosity of Venus.
CHAPTER TEN: APPLICATIONS OF THE MODEL TO HISTORY OF SCIENCE

On the basis of such considerations Palter concludes that Copernicus’s system is not, in its historical context, perceptibly superior to Ptolemy’s in predictive accuracy. In order to square this fact with the putative reality of a “Copernican revolution”, according to Palter, ‘one is constrained to fall back on the criterion of simplicity.’

That the chief virtue of the Copernican theory was the degree of its simplicity – and, frequently, that this was the theory’s sole virtue – has been suggested by many who have surveyed the period. Historians’ attempts to decide which was the relative simplicity of the Ptolemaic and Copernican theories have almost invariably been based on a straightforward count of the numbers of circles postulated by the geometrical constructions to which the two theories appeal. Two figures issue from most such discussions: historians have calculated that Ptolemy’s or a later formulation of the geocentric theory has need of about 80 circles, whereas the Copernican theory is reputed to appeal to some 34 only.

If computation of these figures exhausted the question of the comparative simplicity of the two theories, little doubt could remain about the great improvement effected by the later theory over the earlier. But the issue is surely more intricate than this: the procedure of counting circles is insufficient to gauge the degree of simplicity of an astronomical theory. While it is plausible to assume that the number of circles to which an astronomical theory is compelled to appeal is a factor contributing to determine the overall simplicity of that theory, it is wrong to believe that this number on its own constitutes a measure of that simplicity. Palter makes the reasonable point that ‘comparative counts of circles in astronomical systems can be taken as significant measures of simplicity only when certain definite conditions are satisfied by these systems.’ He suggests a more general view of what constitutes simplicity in an astronomical system: ‘It is the number of independent parameters rather than the geometrical details which

15 Palter (1970), p. 114. Palter denies also the suggestion that the Copernican theory was superior to the Ptolemaic in physical plausibility.
16 Ibid., pp. 114-5.
17 For instance Reichenbach writes: ‘Copernicus [...] was able, in fact, to cite as a distinct advantage only the greater simplicity of his system’ (1927, p. 18).
18 This is roughly the view of e.g. Kordig (1971, p. 109), who states that Copernicus simplified Ptolemaic astronomy by reducing the number of epicycles ‘from 84 to about 30’. For further details and examples of the count see Palter (1970), pp. 94, 113-4, or Cohen (1985), p. 119. The strictly analytical aspects (as distinct from its historiographical aspects) of the question of determining the degree of simplicity of the Copernican system have been addressed by Rosenkrantz (1977), pp. 135-61.
is significant.\textsuperscript{20} On this advice, discussions of relative simplicity ought to focus on
the numbers of parameters of which the values require independent specification
in order for each of the theories to be constructed or applied to a problem in
planetary astronomy.

In attempting to determine the relative simplicity of the two theories on
this generalized basis, an important difference between the aims of Ptolemaic and
Copernican astronomical practice comes to light. The Ptolemaic theory did not
aspire to constitute a systematic model of the heavens: it consisted of a set of
logically independent schemes each of which aimed to describe the motions of
one of the seven known celestial bodies. True, in the scheme dedicated to each
body there appeared a point which, in consequence of its definition, was common
to all the schemes: this point is the centre of the deferent, at which was situated
the earth. But it lay beyond the intentions of the Ptolemaic astronomers that these
schemes should be superposed at this point to yield an exhaustive and systematic
model of the heavens. To this feature of Ptolemaic astronomy there correspond an
ontological disadvantage and a methodological advantage. The disadvantage is
that the Ptolemaic theory fails to constitute a truly unified model of the heavens;
the advantage is that the solution to a problem of observational planetary
astronomy requires less than the full panoply of Ptolemaic constructions, and is
hence attained more simply than if the theory possessed greater
interconnectedness. No single calculation confronting the Ptolemaic astronomer
required the use of all 80 epicycles; in fact, no single planetary problem ever
necessitated appeal to more than the six epicycles governing the motions of the
planet to which the problem referred. Copernicus's astronomy on the contrary
was intended as and amounted to an all-embracing model of the planetary
system. This difference had practical repercussions: calculations of planetary
positions on the Copernican model possessed an interconnectedness not shared by
the analogous computations on the Ptolemaic model. A Copernican could not
compute the apparent position of Mars at some time without reference to the
position of the earth in its orbit at that instant.\textsuperscript{21} In this sense, as a set of solutions
to individual problems, Ptolemaic astronomy is much simpler and convenient – if

\textsuperscript{20} Ibid., p. 97.

\textsuperscript{21} In this spirit Price (1959, p. 199) identifies the originality of Copernicus's
contribution in the construction of a mathematical planetary system, as distinct from
a mathematical theory of the individual planets. Further on this difference between
Ptolemaic and Copernican astronomy see Hanson (1961), pp. 175-7. In view of the
fact that Ptolemaic astronomy stops short of constituting a system, Kuhn's
customary use of the expression 'the Ptolemaic system' (1957, e.g. p. 71) is
somewhat misleading.
rather less systematic – than the Copernican system.

There is good evidence for believing that the mature Copernicus realized that he could claim on behalf of his system a degree of simplicity no greater than that of the Ptolemaic theory. His early work, the Commentariolus, had referred to the degree of simplicity of the Copernican system as indicating its superiority to the Ptolemaic theory. If Copernicus had persisted in the belief that his theory was superior to its competitor because it was simpler, he would surely have repeated and elaborated on the claim in his more systematic treatise, the De revolutionibus, in the same way in which most of the other claims in favour of the Copernican theory which appear in the Commentariolus receive an extended treatment in the later work. Instead, the De revolutionibus omits to claim the degree of simplicity of the Copernican theory as a respect in which it demonstrated superiority to the Ptolemaic alternative. In fact, the claims to simplicity which figure in the Commentariolus appear to be replaced in the later publication by claims that the Copernican theory was preferable to the Ptolemaic on the strength of its superior internal harmony. One concludes from such observations that at the time of its enunciation the Copernican theory was neither simpler nor reputed to be simpler than the Ptolemaic theory.

Both predictive accuracy and degree of simplicity are thus revealed as features in which the Copernican theory could not have hoped to demonstrate superiority over the Ptolemaic theory. Kuhn agrees that whatever superiority the Copernican theory may have demonstrated, it could not have been a matter of predictive accuracy or simplicity: 'Judged on purely practical grounds, Copernicus' new planetary system was a failure; it was neither more accurate nor significantly simpler than its Ptolemaic predecessors.'

It does not appear from the above that the Copernican theory either could have legitimately acquired or in fact did acquire adherents in its community on the strength of its possession of the two logico-empirical features, its degree of predictive accuracy and the degree of its simplicity, or at least in virtue of its

---

22 For the passage of the Commentariolus which advances on behalf of the Copernican system the claim of simplicity see Swerdlow ed. (1973), pp. 434-6.
23 The shift of Copernicus's claim from the supposed simplicity of the geocentric system, as in the Commentariolus, to its internal harmony, as in the De revolutionibus, is further discussed in Pera (1981), pp. 157-9. The significance of the De revolutionibus's claim to internal harmony is discussed below.
24 Additional grounds in support of this claim have been adduced by Cohen (1960), p. 58, Neugebauer (1968), and Lakatos and Zahar (1975), pp. 360-4. Gingerich (1975, p. 87) concludes that 'the Copernican system is slightly more complicated than the original Ptolemaic system.'
CHAPTER TEN: APPLICATIONS OF THE MODEL TO HISTORY OF SCIENCE

possession of these features to a degree markedly greater than that of the Ptolemaic theory. The present treatment will proceed to suggest that the grounds upon which the Copernican theory won adherents from the Ptolemaic were aesthetic.

From the time of classical Greek astronomy to the sixteenth century, planetary astronomy was founded on two principles: the principle of geocentricity, and the principle of uniform circular celestial motions. Copernicus wrought a break with one component of this tradition in repudiating the commitment of planetary astronomy to the principle of geocentricity but maintaining that to the principle of circularity: he retained the notion that the planets move uniformly in orbits that are circular, or compounded of Ptolemaic epicycles. The principle of geocentricity had been rejected by many astronomers before Copernicus, as the De revolutionibus itself took pains to stress in order to reduce the apparent novelty of the work. In relaxing adherence to this principle Copernicus thus followed a relatively familiar methodological path. The commitment to the principle of circularity had on the contrary not been weakened for 2,000 years. The reason for this was in part its entrenchment in Aristotelian natural philosophy. Aristotle had allotted celestial bodies to the superlunary sphere; because of their location, only a certain and perfect type of motion could legitimately be attributed to them. Circular motion possessed the required perfection in virtue of its following two characteristics. First, circular motion does not admit of an end-point, and hence is on the Aristotelian view the motion of a self-contained or perfect mover and not motion towards that which the mover lacks; second, motion of a sphere in a circle is the only form of motion in which the mover remains forever superposed to its own previous positions. Motion in a circle thus pertains to the very essence of a celestial body.

Copernicus signals his acceptance of this Aristotelian view:

It is impossible for a heavenly body which is simple to move irregularly in a single sphere. That would have to be due either to changes in the moving power, whether derived from elsewhere or from its intrinsic nature, or on account of unevenness in the revolving body. Both these possibilities are unacceptable to the reason, and it is inappropriate to attribute such a thing to bodies which are established in an ideal state.


Copernicus appeared to find it an inderogable requirement of planetary theory that it should subscribe to the principle of the circularity of celestial motions. He attacked Ptolemaic astronomy in both the *Commentariolus* and the *De revolutionibus* not as a heliocentric theorist criticizing a geocentric theory, but because Ptolemy had adhered insufficiently strictly to the precept that all celestial motions must be explained only by uniform circular motions or combinations of such motions. Ptolemaic astronomers had surmised that, within the constraints of geocentrism, accuracy of representation of planetary motions necessitated the relaxation of this precept, and the attribution to planets of non-uniform circular motions, the fluctuations of which are determined by the location of the so-called *punctum aequans.* The introduction of the equant point permitted the most accurate representations of planetary motions that would be attained by Western astronomy until the time of Kepler. Copernicus however considered appeal to equants a violation of a fundamental principle and devised a system in which celestial bodies moved in uniform motion along circles or combinations of circles. This he achieved in the *De revolutionibus.*

Copernicus seemed to believe that the restoration of the principle of uniform circular celestial motions was not only a notable achievement in planetary astronomy, but also the feature of his theory which indicated its superiority to its Ptolemaic competitor. In the intentions of Copernicus, the more faithful adherence by his theory than by that of Ptolemy to the principle of uniform circular motions was a feature of the sort required to encourage the transfer of allegiance from the Ptolemaic theory towards his own.

There is evidence to suggest that indeed it was this feature of Copernican theory rather than any other which attracted adherents in his community. For instance, the favourable opinion of the work of Copernicus held by one of the leading astronomers of his time, E. Reinhold, appears to have been motivated more by its elimination of the equant and its return to pure uniform circular motion than by its substitution the sun for the earth at the centre of the universe.

Now, what sort of theoretical feature is the quality of adhering to the principle of uniform circular celestial motions? Chapter 5 subdivided features of theories into the two classes of empirical and non-empirical, or aesthetic.

---

29 The *punctum aequans* was a point positioned so that the radiant vector centered upon it which traces the motion of the planet in its orbit rotates with uniform angular velocity.

Empirical features of theories were those deemed constitutive of the attainment of the goal of science, observational success. The quality of adhering to the principle of uniform circular celestial motions is not constitutive of observational success: the degree of observational success attained by theories in planetary astronomy appears not to be correlated to the strength with which they affirmed this principle.\(^{31}\) If one accepts the bipartition advanced in the present treatment, this suggests that the feature of Copernican theory which contributed most to winning adherents to its side was aesthetic in nature.

Copernicus's expectation that his theory would win support on its aesthetic virtues is visible in the *De revolutionibus* itself. Its text claims as the chief merit of the theory a quasi-aesthetic internal harmony greater than that of the Ptolemaic theory:

> Those who have devised eccentric circles [...] have not been able to discover or deduce from them the chief thing, that is the form of the universe, and the clear symmetry of its parts. They are just like someone including in a picture hands, feet, head, and other limbs from different places, well painted indeed, but not modelled from the same body, and not in the least matching each other, so that a monster would be produced rather than a man. \(^{32}\)

Kuhn too perceives an aesthetic concern in the arguments deployed in the *De revolutionibus*:

> Each argument cites an aspect of the appearances that can be explained by either the Ptolemaic or the Copernican system, and each then proceeds to point out how much more harmonious, coherent, and natural the Copernican explanation is. [...] Copernicus' arguments are not pragmatic. They appeal, if at all, not to the utilitarian sense of the practicing astronomer but to his aesthetic sense and to that alone. [...] The harmonies to which Copernicus' arguments pointed did not enable the astronomer to perform his job better. New harmonies did not increase accuracy or simplicity. Therefore they could and did appeal primarily to that limited and perhaps irrational subgroup of mathematical astronomers whose Neoplatonic ear for mathematical harmonies could not be obstructed by page after page of complex mathematics leading finally to numerical predictions scarcely better than those which they had known before.\(^{33}\)

---

\(^{31}\) In point of fact, of course, the greater observational success in sixteenth- and seventeenth-century planetary theory was attained by a theory which violated the principle of uniform circular celestial motions, the theory of Kepler.

\(^{32}\) Copernicus (1543), p. 25 (Prefatory letter to Pope Paul III). Copernicus returns to praise the internal harmony displayed by his theory later in the same work (*ibid.* p. 50, Book I, chapter 10): 'We find [...] in this arrangement the marvellous symmetry of the universe, and a sure linking together in harmony of the motion and size of the spheres, such as could be perceived in no other way.' For yet another similar formulation see *ibid.*, p. 233 (Book V, preface).

\(^{33}\) Kuhn (1957), p. 181; emphasis in the original.
Neugebauer appears to agree that the entrenchment of Copernicanism was achieved not through logico-empirical strengths of the theory such as its degree of simplicity, but through its aesthetic virtues, or the degree to which it appeared ‘pleasing to the mind’: ‘Had it not been for Tycho Brahe and Kepler, the Copernican system would have contributed to the perpetuation of the Ptolemaic system in a slightly more complicated form but more pleasing to philosophical minds.’

On such a basis one concludes that the grounds upon which the Copernican theory gained endorsement in the period after its first formulation were aesthetic rather then logico-empirical in nature.

If a theory garnered adherents on the strength of its aesthetic features rather than or as well as in virtue of its logico-empirical qualities, it must have constituted according to the present treatment a non-revolutionary innovation: for only an innovation which effected with past science a break less profound than a revolution could hope to find its own aesthetic features valued by the pre-existing aesthetic evaluative canons of the community into which it is received. According to the historiographic findings mustered above, the Copernican theory maintained aesthetic continuity with previous theories in its discipline, and thus cannot have constituted a revolutionary innovation. On the present treatment the Copernican theory amounts merely to a non-revolutionary – albeit perhaps very important – innovation in the fabric of mathematical planetary astronomy. The return to the Greek canons of circularity and uniformity could not have constituted a revolution except in the sense of a return to the ideals of the past.

The conclusion that Copernicanism constituted something less than a revolutionary innovation in planetary astronomy has been reached here by the application of a definite model of scientific revolutions, which suggests inter alia that the question whether or not a certain theory amounted to a revolutionary break with past science can be decided by an examination of the grounds on which that theory gained adherents in its community. This conclusion is nonetheless shared by several historians of science who, setting aside the received

---

34 Neugebauer (1968), p. 103. The remainder of that paper amounts to a masterly study of the dependence of Copernican theory upon Ptolemaic. Contributors to Neyman ed. (1974) also examine the extent to which Copernicus constructed a theory which was ‘pleasing to the mind.’

35 The aesthetic features of Copernican theory which may have encouraged its adoption have been studied by Gingerich (1975, pp. 89-90) and by Hutchinson (1987, pp. 109-36). In the work cited here Gingerich presents the Copernican theory as a resolution of a state of aesthetic dissatisfaction which had been accumulating with the development of Ptolemaic astronomy.

36 This thesis is argued also by Cohen (1985), pp. 123-5.
belief in the occurrence of a ‘Copernican revolution’, have investigated afresh the nature of the transition from Ptolemaic to Copernican astronomy. Since the judgements of these historians were reached on grounds different from those adduced in the present treatment and by arguments which do not depend on the application of the present model of scientific revolutions, their conclusions offer an independent check of the historical verisimilitude of this model.

Broadly speaking, two series of historiographic studies have tended to suggest that Copernican theory maintained with pre-existing astronomy a degree of continuity too great for that contribution to be termed ‘revolutionary’ without a severe dilution of the concept. Studies of the first series have examined internal features of Copernican theory for evidence of its intellectual derivation from or continuity with Ptolemaic astronomy: they have remarked on such matters as the Ptolemaic inspiration of the text of De revolutionibus or the Aristotelianism of the doctrine of the spheres to which Copernicus appeals. In both these respects as in several others the conclusions appear to endorse Hanson’s characterization of the De revolutionibus: ‘It was a comprehensive attempt to make the science of that day work better; it was not explicitly a plan for a new science of tomorrow.’

Studies of the second series have investigated the reception accorded to Copernican theory to gauge the degree to which the theory was perceived as ground-breaking or unorthodox by the community of mathematical planetary astronomers. R.S. Westman concludes that, far from being perceived as iconoclastic, the Copernican theory was respectfully welcomed into what Kuhn would term the ‘normal science’ of Ptolemaic astronomy. Westman points out that if Copernicanism truly effected a revolutionary break with Ptolemaic theory and constituted a new paradigm in mathematical planetary astronomy, then his historiographic findings about the manner of the reception of this theory in the community contradicts Kuhn’s view of the genesis of new paradigms at times of disciplinary crisis: Kuhn had after all predicted that new paradigms would be

---

37 One of the earliest consequential histories of astronomy to suggest that the work of Copernicus is best interpreted as the culmination of a long-standing project is that of Pedersen and Pihl (1974), pp. 299-314.

38 The striking textual similarities between the presentation of Copernican theory in the De revolutionibus and the exposition of Ptolemaic astronomy in the Almagest have been noted by Neugebauer (1952, p. 206): ‘chapter by chapter, theorem by theorem, table by table, these works run parallel.’ One may from this formal resemblance seek to infer that Ptolemaic and Copernican astronomy shared elements of the same methodology, or intended to confront the same complex of problems. Copernicus’s doctrine of the spheres and its resemblance to Ptolemy’s is examined in Jardine (1982).


whole-heartedly either embraced or resisted, but would in no case be susceptible to accommodation within the framework of the pre-existing paradigm. Westman thus suggests that the transition from Ptolemaic to Copernican astronomy may constitute an unfavourable test-case for Kuhn’s model of scientific revolutions. The view taken on this matter by the present treatment is different from Westman’s: on the present account the Ptolemaic-Copernican transition fails to constitute a counterexample to Kuhn’s theory of the manner in which revolutions unfold, but merely because that transition did not constitute a revolution and hence is immaterial to any test of this component of Kuhn’s model.41

3. Aspects of revolution: the Keplerian ellipses

The previous section suggested that Copernican theory did not represent a revolutionary innovation in mathematical planetary astronomy, and that this is shown by the fact that it was able to gather support in its community on the strength of its aesthetic features. This section will proceed to consider the second transition in early-modern astronomy, that wrought by Kepler’s theory of planetary motions.42 This treatment will investigate the identity of the grounds on the strength of which Kepler’s theory was able to garner support, and surmise from such historiographic data the nature of the innovation which that theory constituted.

Enunciations of Kepler’s first two laws of planetary motion are contained in his Astronomia nova of 1609. They were the fruit of the author’s ‘war on Mars’, his effort between 1600 and 1605 to discover a law describing the motions of the sun’s fourth planet.43 This work was an prolonged attempt to formulate a mathematical description of the orbit and motions of Mars which would demonstrate consistency with the data collated from observations of the positions of the planet by Tycho Brahe. These data boasted an accuracy of around 1": they were substantially more accurate than any similar data previously collected, and

41 Of course the present treatment contradicts another component of Kuhn’s overall view of the history of science, viz., his belief – expressed in both his (1957) and (1962) – that the Copernican theory constituted a revolution.
42 This treatment will consider not the doctrines contained in Kepler’s Mysterium cosmographicum of 1596, but those of the Astronomia nova of 1609 and his later works.
43 The chains of reasoning which led Kepler to his first law of planetary motion are retraced in Whiteside (1974); those which led him to his second law in Alton (1969).
their precision would not be bettered until the advent of the telescope. Kepler appears to have attained his first law – the proposition that the planets move in an ellipse with the sun at one focus – by, roughly speaking, an alternation of theoretical hypothesis and empirical test: he proposed a succession of candidate-paths for the orbit of Mars and gauged the accuracy with which the quantitative implications of each hypothesis cohered with Tycho’s data.

Kepler tested at the outset the hypothesis, typical of strict Copernicanism, that Mars moved in a circular orbit. The trajectory which in this event Mars would have traced departed from the planet’s observed path by an angular distance of up to 8'. This discrepancy, the magnitude of which is much greater than the margins of error claimed for Tycho’s data, was in Kepler’s view sufficiently large for a circular orbit to be ruled out.44

The distribution of the discrepancies between the observed path of Mars and its path on a putative circular orbit suggested to Kepler in 1602 the curve which he should next consider: The orbit is not a circle, but [passing from aphelion] enters in a little on either side [at quadratures] and goes out again to the breadth of the circle at perihelion, in a path of the sort called an oval.45

However the putative oval orbit too diverged from from the true orbit as this had been established from observations. The sign of the crucial divergences exhibited by the oval orbit was opposite to that of the divergences introduced by the circular orbit. Kepler concluded from this in 1604 that the path of the true orbit was a curve contained between the circle and the oval, and in the same breath suggested which curve this might be:

In the middle longitudes […] the perfect circle prolongs [the true orbital path] by about 800 or 900 [parts in 152350, the mean radius of orbit] too much. My ovality curtails by about 400 too much. The truth is in the middle, though nearer to my ovality […] just as though Mars’s path were a perfect ellipse.46

Kepler found that the hypothesis that the orbit of Mars was an ellipse bearing the sun at one of the foci led to a close accord with the data. The content of the first law of planetary motion which Kepler published in the Astronomia nova expressed this discovery.

The role played in this research by logico-empirical evaluative criteria is obvious.47 Kepler had set himself the problem of identifying a closed curve which

---

44 This passage in Kepler’s reasoning is described in Whiteside (1974), pp. 6-7.
45 Cited ibid., p. 8; interpolations by Whiteside.
46 Cited ibid., p. 11.
47 The empirical content of Kepler’s laws is analyzed by Mittelstrass (1972), pp. 208-9.
was consistent with a set of numerical data; after having rejected at least two
possible curves on the strength of their divergence from observational data, he
found the hypothesis that Mars traced an elliptical orbit to be the only one which
demonstrated a sufficiently close accord with the data. Clearly thus Kepler's
theory had upon publication in 1609 already undergone evaluation on logico-
empirical criteria; indeed the theory which he published in that year was the sole
survivor, out of a pool of several candidates, of a preliminary bout of theory-
evaluation on logico-empirical criteria.

Here is some indication of the nature of the grounds on which Kepler's
theory gained support upon its formulation, even if only in the mind of its
begetter. More substantial evidence of the grounds upon which the theory gained
supporters in its community is drawn from the story of its later reception.\(^48\)

The greatest empirical triumph of the Keplerian theory of the years
immediately following its enunciation, and the achievement which more than any
other contributed to its establishment in its community, came with the publication
by Kepler in 1627 of the *Tabulae Rudolphinae*. These were compilations of
numerical predictions of the future positions of the moon and planets computed
by the application of Kepler's laws, and which offered themselves with the
passage of time to testing against observational data by planetary astronomers. In
essence these tables were an example of an entity frequently discussed by
philosophers of science but seldom so concretely and explicitly produced in
scientific practice: they were a tabulation of the observational consequences of a
body of theory, which by their means opened itself to experimental test. The
project of testing Keplerian theory was facilitated by the publication of the
Rudolphine tables more than one might have expected, for astronomers of the
time of Kepler had much less familiarity with the properties of the ellipse than
with those of the circle, and had -- needless to report -- no recourse to coordinate
geometry or differential calculus; they would hence have found the task of
computing the predictive consequences of the theory ungrateful. From the
moment of the publication of the tables, astronomers acquired means which
enabled them easily to test the predictions of Keplerian theory against the actually
observed positions of the sun, moon and planets, and compare the accuracy
attained with that of rival astronomical theories.\(^49\) Use of the Rudolphine tables

---

\(^{48}\) The reception gained by Kepler's theory in the period 1609-66 is described in J.L.
Russell (1964).

\(^{49}\) On the contribution of the Rudolphine tables to the empirical testing of Kepler's
quickly showed that the ellipse which Kepler had successfully posited as the form of the orbit of Mars fitted with good precision also the orbits of the earth and the other planets, even that of Mercury which had thus far proved the planetary orbit most recalcitrant to astronomical theories.\textsuperscript{50}

This historical evidence suggests in a preliminary way that the grounds upon which Kepler's theory established itself in the community of mathematical planetary astronomers were chiefly logico-empirical in nature: it was the demonstrated close accord of the Keplerian predictions in the Rudolphine tables with observation which played the largest role in convincing members of the community that they ought to embrace Kepler's theory in preference to the Ptolemaic, the Tychonic or the Copernican.

The model of scientific revolutions contained in chapter 8 suggests that a theory which constituted a revolutionary innovation in its discipline will gain adherents exclusively on the strength of its logico-empirical virtues, remaining unable to gather support in virtue of its aesthetic features in view of its aesthetic innovativeness. Thus, the historiographic finding that Keplerian theory attracted adherents on the strength of its empirical predictions is at least consistent with the thesis of this section, that Kepler's theory constituted a revolutionary innovation in the discipline of mathematical planetary astronomy. More detailed historiographic research offers stronger support for this thesis, as will be shown next.

The model of scientific revolutions constructed in the present treatment suggests that the reception of a theory which constitutes a revolutionary innovation will exhibit two phases. The first will be characterized by resistance to the new theory, especially on the part of the more conservative members of the community, on the grounds of its aesthetic innovativeness, which causes its aesthetic features to conflict with the aesthetic evaluative canons to which part of the community remains wedded. A theory which constitutes a successful revolution will notwithstanding this resistance proceed to demonstrate substantial empirical success, and in particular greater such success than can be mustered by the preceding, aesthetically conservative theory which constitutes its competition. The second distinctive phase in the reception of such a theory will be inaugurated when the empirical success accumulated by the new theory is so substantial as to cause the greater part of the community – including many members who were previously opposed to it on aesthetic grounds – to embrace it in order to draw

\textsuperscript{50} As Russell (1964, p. 20) recounts, Keplerian astronomy attracted far more favourable attention after the publication of the Rudolphine tables than before it.
benefit from its empirical power, now demonstrably greater than that of its competitor-theory. To further this end, members of the community will suspend or de-emphasize their aesthetic sensibility towards theories in order to avoid the prescriptive conflict which would result from an application at full strength of the two evaluative canons.

Each of these distinctive phases of revolution was realized in the community which received Kepler’s theory. Kepler’s assertion that planets moved in elliptical orbits encountered a system of astronomical belief unfavourably disposed towards the suggestion that celestial bodies might move in paths other than circular. This hostility was justified by those who professed it not on the empirical criterion that the postulation of non-circular orbits rendered planetary theory incapable of accounting with sufficient accuracy for astronomical data, but on a non-empirical, metaphysical and – on the present treatment – aesthetic criterion, that only the postulation of circular orbits could attain the harmony demanded for planetary astronomy. This attitude, widely held in the community in which Kepler worked, was manifested for instance by Tycho in a seemingly admonitory passage of a letter to Kepler:

The orbits of the planets must be constructed exclusively from circular motions; otherwise they could not recur with a uniform and equal constancy, eternal duration would be impossible; moreover, the orbits would be less simple, would exhibit greater irregularities and would not be suitable for scientific treatment and practice.\textsuperscript{51}

In this passage Tycho gives notice of opposition to the assertion which was to become the distinctive thesis of Kepler’s theory. His objection is moved on archetypically extra-empirical or more precisely – in the eyes of the present treatment – aesthetic grounds. Tycho resists the introduction of non-circular orbits into planetary theory not because the observational success of the theory which posited such orbits would be impaired, but because the resulting theory would be less attractive on aesthetic, metaphysical grounds. This stance is characteristic of the conservative faction of the community in which a theory which constitutes a revolutionary innovation has been proposed: as the present treatment predicts, members of that faction resist the new theory not in virtue of the application of the logico-empirical criteria of theory-assessment – which may yield a positive evaluation of the theory’s predictive power – but on the grounds of the new theory’s violation of the pre-existing aesthetic evaluative canons of the

community. Tycho’s statement thus expresses the attitude which the present model predicts will be widespread in the first phase of the response to a revolutionary theory.

The present treatment further predicts that as the new theory demonstrates its possession of logico-empirical qualities to a degree superior to that of previous theories which adhered to the extant aesthetic canons, then its violation of those canons will come to assume less and less importance for theory-choice in the eyes of the community. In time even the conservative faction will be led to suspend its faculties of aesthetic theory-evaluation and embrace the new theory on logico-empirical grounds. Of this phenomenon too the history of the reception of the Keplerian theory offers evidence. In the years subsequent to the publication of the *Astronomia nova*, P. Crüger, Professor of Mathematics at Danzig, expressed profound scepticism towards Keplerian astronomy and recoiled from entertaining it. He wrote for instance in 1624: ‘I do not subscribe to the hypotheses of Kepler. I trust that God will grant us some other way of arriving at the true theory of Mars.’\(^{52}\) The publication in 1927 of the *Tabulae Rudolphinae* caused him radically to change his mind. In a letter to the astronomer P. Müller of 1629 – thus after the tables had had the opportunity of showing the predictive accuracy of the theory on which they were based – Crüger expressed the impact which the empirical corroborations had caused upon him:

> You hope that someone will give these tables [the astronomical tables of Longomontanus] a further polishing and you say that all astronomers would be grateful for this. But I should have thought that it would be a waste of time now that the Rudolphine Tables have been published, since all astronomers will undoubtedly use these. [...] I am wholly occupied with trying to understand the foundations upon which the Rudolphine rules and tables are based, and I am using for this purpose the Epitome of Astronomy previously published by Kepler as an introduction to the tables. This epitome which previously I had [...] so many times thrown aside, I now take up again and study [...]. I am no longer repelled by the elliptical form of the planetary orbits [...].\(^{53}\)

In the last sentence of this passage, Crüger renounces one of the criteria upon which he had at a previous time unfavourably assessed Keplerian theory: he withdraws from among the grounds upon which to oppose the new theory the observation that it attributes to planetary orbits an elliptical form. The feature of Kepler's theory which Crüger no longer wishes to count against it is non-

\(\text{\textsuperscript{52}}\) Here cited in the translation of Russell (1964), p. 8. For further expressions of Crüger's early unfavourable response to Kepler's theory see *ibid.*, pp. 7-8.

\(\text{\textsuperscript{53}}\) Cited *ibid.*, p. 8.
empirical or aesthetic in nature: it is the theory’s violation of the pre-existing
metaphysically-backed prescription of a certain form of simplicity or symmetry to
planetary theory, to which Ptolemaic, Tychonic and Copernican theory had each
adhered.

The reason for which Criiger feels he can no longer afford to reject
Keplerian theory on the grounds of its violation of this principle is, as the
remainder of the above passage makes clear, that the theory had manifested
through the Rudolphine tables a high degree of empirical accuracy. The present
model of scientific revolutions predicts that at least some members of the
conservative faction of a scientific community – those who had initially
condemned a revolutionary theory on the grounds of its infraction of aesthetic
canos – will later be drawn nonetheless to embrace it in recognition of the great
degree to which it satisfies the logico-empirical criteria of theory-evaluation. It is
to the credit of Criiger if he was able after his study of the evidence contained in
the Rudolphine tables to set aside his initial non-empirical reservations against
Keplerian theory and acknowledge that its empirical power rendered its adoption
advisable.

The grounds adduced in evaluations of Kepler’s theory during the period
of its first reception enable conclusions to be drawn – in accordance with section
1 of this chapter – about the nature of the innovation constituted by that theory
in planetary astronomy. Unlike Copernicus’s theory, which the previous section
showed to have gained support on the strength of its aesthetic features rather
than of any supposed significant empirical superiority over Ptolemaic theory,
Kepler’s theory appears to have attracted endorsement in virtue of the
considerably greater accuracy of its empirical predictions over that of any
previous theory, and in spite of its non-empirical or aesthetic features. Far from
reinforcing the appeal of the theory among Kepler’s contemporaries, its aesthetic
features proved a hindrance to the theory’s establishment which had gradually to
be overcome by demonstrations of its empirical power. Unlike Copernicus’s
theory, Kepler’s thus meets the specifications which the present treatment
prescribes of a theoretical innovation for it to be considered revolutionary: it
appears to have effected a rupture with the pre-existing aesthetic evaluative
canos of its community, and it established itself in the teeth of initial aesthetic
opposition on the strength of the considerable degree to which it exhibited the
possession of logico-empirical virtues.\textsuperscript{54}

Kepler's theory appears on these grounds the crucial turning-point between medieval planetary astronomy and Newtonian celestial dynamics. The present treatment thus agrees with the verdict on this period of Hanson: 'The line between Ptolemy and Copernicus is unbroken. The line between Copernicus and Newton is discontinuous, welded only by the mighty innovations of Kepler.'\textsuperscript{55} Certainly Kepler was responsible for more than merely a 'version of Copernicus' proposal', as Kuhn seeks to construe his contribution.\textsuperscript{56}

The last two sections have attempted to support the model of scientific revolutions presented earlier in this treatment in the following way. They have corroborated on historiographic grounds the model's description of a community's reactions to theories which constitute revolutionary and non-revolutionary innovations in their discipline. To this extent the story of the transition from Ptolemaic-Copernican to Keplerian astronomy demonstrates the application of aesthetic canons of theory-appraisal in a phase of scientific history, and hence permits one to gauge the ability of the model in its entirety to account for historiographic data.\textsuperscript{57}

\section*{4. Aspects of continuity: the theory of special relativity}

The discipline of physics between 1900 and 1940 underwent transformations perhaps even more conspicuous than those of astronomy between 1540 and 1630.

\textsuperscript{54} That Kepler's theory was responsible for a revolutionary break has been affirmed by historians on arguments entirely independent of those pursued in the present treatment of scientific revolutions: see e.g. Mittelstrass (1972), pp. 205, 207.

\textsuperscript{55} Hanson (1961), p. 169.

\textsuperscript{56} Kuhn (1957), p. 219.

\textsuperscript{57} There is an aspect of seventeenth-century astronomy separate from those pursued in this chapter which allows the incidence of aesthetic factors in theory-choice to be studied in abstraction from empirical grounds. Throughout the seventeenth century the community of mathematical planetary astronomy possessed two theories - Copernican heliocentrism and Tychonic geo-heliocentrism - which were observationally precisely equivalent. Any choice between them could therefore be made solely on non-empirical grounds. Undoubtedly among the factors which decided the choice for some natural philosophers were arguments inspired from religious doctrine and physical plausibility; but a study of the preferences of mathematical scientists of this period between the two equivalent theories could reveal also a vein of aesthetic motivations and argumentation. I propose to address this question in forthcoming research.
The theories of relativity and of quantum mechanics vie with one another in popular accounts for the title of the supreme innovation of modern science.

The task of discerning which, if either, of these contributions to physics constituted a revolution will – as is by now usual in this treatment – be discharged by means of an examination of the grounds on which each gained adherents early after their first enunciations. As has been repeated above, the present model prescribes that the discovery that a new theory gained endorsement on aesthetic grounds indicates that it failed to constitute a revolution, whereas the finding that it attracted support on logico-empirical grounds and in the face of aesthetically-motivated opposition suggests that it indeed constituted a revolutionary departure from previous science.

The old suggestion that Einstein was impelled to formulate the theory of special relativity by the null results of the aether-drift experiments of A.A. Michelson and E.W. Morley has by now been conclusively discredited. Far from playing an important role in convincing Einstein of the truth of his theory as he worked on it, these results came to his notice only some time after the publication of his paper on special relativity of 1905. In fact, the paper does not invoke explicitly any of the experimental results unfavourable to elements of classical physics, and the Michelson-Morley experiments are not mentioned even when the opportunity arises to show how the theory of relativity would account for their result. It is thus illegitimate to attribute the entrenchment of the theory of relativity in Einstein’s mind in the period in which he conceived it and refined its formulation to this empirical virtue of the theory, its ability – not shared by classical physics – easily to account for the Michelson-Morley results. What were then the grounds upon which Einstein, even as he formulated it, was first persuaded of the value of the theory of relativity?

In view of the fact that theory-assessment consists generally of a comparative evaluation of two or more alternative theories rather than of an appraisal of an isolated theory, to answer this question one must first investigate the reasons for Einstein’s dissatisfaction with classical physics, the theoretical corpus constituted essentially by the conjunction of Newtonian mechanics and Maxwell’s electrodynamics. Einstein indicates the cause of his dissatisfaction in

---

CHAPTER TEN: APPLICATIONS OF THE MODEL TO HISTORY OF SCIENCE

the opening of the paper in which the theory of relativity is first expounded: 'That Maxwell's electrodynamics - the way in which it is usually understood - when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena is well known.'

The asymmetries of which Einstein here thinks appear in the application of Maxwell's electrodynamics to certain physical systems, including the system composed of a conductor and a magnet in motion relative to one another. In, at least, their formulations cognizant of the work of Mach, the principles of classical physics recognize relative motions but none that is absolute. One is hence led by classical physics to expect no difference between the situation obtained by moving a conductor at a certain velocity relative to a magnet which is at rest in the laboratory and that produced by now affixing the conductor to the bench and moving the magnet with the same velocity respect to it. It is expected that all physical parameters will take identical values in the two cases: in particular, the intensities of the electrical currents induced in the conductor are expected to be identical in the two cases. Experiment bears out these expectations.

Maxwell's electrodynamics predicts to high accuracy the intensity of the induced currents for any values of the velocities of conductor and magnet. Of greater interest to Einstein than the outcome of the calculations was however their route. Maxwell's electrodynamics prescribes that one apply to the system composed of stationary conductor and moving magnet an analysis substantially different from that which it deems appropriate to the system of stationary magnet and moving conductor. No matter that the final predictions of these analyses coincide both with one another and with empirical data: Einstein found displeasing the unwarranted asymmetry which was embedded in the theory.

The theory of special relativity was, on Einstein's own account, designed to provide a treatment of systems of conductors and magnets in relative motion which exhibited symmetry in the sense specified above, and therefore did not offer analyses of these systems which differed according to whether the motion

---


63 For a study of the way in which classical electrodynamics treats the system of magnet and conductor see Miller (1981), pp. 145-50.

64 As Holton (1973, p. 168) points out, Einstein's displeasure at the asymmetries of classical physics is voiced not only in his paper on special relativity but also in his other two important papers of 1905. Further on Einstein's concern for symmetry as a property of theories see ibid., pp. 362-7. It is only fair to record the contention of Shelton (1988) that Holton overestimates the importance of symmetry considerations as a motivation of Einstein's work.
was attributed to the magnet or the conductor. In this task the new theory succeeded well. That this success was the principle perceived achievement of Einstein's theory of 1905 is intimated by the name by which he intended it to be known: the 'theory of invariants'. These observations suggest that the theory of special relativity gained support at least in part and at least in the system of beliefs of its originator on the strength of the form of symmetry embedded in its structure.

To grounds of what nature does one refer when one favourably evaluates the theory of special relativity for the symmetries of its explanatory structure? In particular, are the grounds which one thus adduces empirical or aesthetic in nature?

Again one must hark back to the grounds for Einstein's dissatisfaction with classical physics. The shortcoming which Einstein perceived in electrodynamics - the asymmetries inherent in its treatment of certain physical systems - was not an empirical defect of the theory: the accuracy of the predictions yielded by Maxwell's electrodynamics for the intensity of the current induced in a conductor in motion with respect to a magnet is both very high relative to the precision of the experimental data and as high as that of the predictions yielded by applications of the same theory to phenomena where no similar asymmetries are manifested. Thus the fact that Maxwell's electrodynamics prescribes that one should choose between two different analyses of a physical system on the basis of the identity of the body which one imagines to be in motion does not appear to impair the predictive power of the theory: positivists - if one indicates by this term scientists concerned solely with the degree of observational success of their theories and not with any non-observational virtues or defects which they may manifest - would find in the asymmetry of applications of Maxwell's electrodynamics no reason to weaken their allegiance to the theory.

The grounds upon which Einstein rejected classical physics provided also the grounds upon which the theory which he proposed gained initial support at least in his own mind: the appeal which the theory of special relativity had for Einstein lay chiefly in its being free of the non-empirical shortcoming which he criticized in classical physics. This quality of special relativity is clearly itself non-empirical: the fact that the new theory omits to follow the old in drawing a distinction between certain physical systems of itself communicates nothing about the predictive virtues of the new theory by comparison with those of its

66 Wider issues in the reception of relativity theory are examined in Glick (1987).
The present treatment suggests that the criterion on which Einstein repudiated classical physics and embraced the theory of relativity was aesthetic in the sense defined in the course of the foregoing chapters. It was on this view an aesthetic dissatisfaction which Einstein felt in classical physics, and on aesthetic grounds that the theory of special relativity proved more satisfactory to him.\(^67\)

The suggestion that Einstein may have made recourse to an aesthetic criterion of theory-assessment in the work leading to the formulation of the theory of relativity is of course supported by the evidence of his methodological beliefs presented elsewhere in this treatment. For instance, chapter 5 construed part of the ‘Autobiographical Notes’ as the description of a two-criterion canon of theory-assessment of which one of the component criteria was aesthetic in nature. Einstein’s choice between classical physics and the theory of relativity is to be seen as an application of this aesthetic criterion of evaluation.

Theories which win adherents to have constituted not revolutionary breaks with previous science but rather moderate innovations, which remain within the paradigm which dominated the discipline at the time of their formulation. The fact that part of the appeal of the theory of relativity was perceived to reside in its aesthetic features leads one to the conclusion that this theory was a moderate, aesthetically-conservative innovation of the latter sort. In maintaining aesthetic solidarity with classical physics, the theory of relativity betrays its membership of the paradigm in physical science which was defined principally by Maxwell’s electrodynamics.

The conclusion that the theory of relativity is most appropriately seen as the culmination of the programme of nineteenth-century physics, here reached on an analysis of the grounds on which it attracted early support and on inferences prompted by the present model of scientific revolutions, is reached by independent itineraries by several philosophically-informed histories of classical physics.\(^68\) Holton for instance writes:

*The so-called scientific ‘revolution’ turns out to be at bottom an effort to return to a classical purity.* [...] Indeed, while it is usually stressed that

---

\(^67\) This view wins the assent of several historians of twentieth-century physics. Swenson (1972), p. 157, concurs that the motivations which led Einstein to the theory of special relativity were largely aesthetic in nature. Pais (1982), pp. 138-40, retracts some of the considerations on aesthetic grounds which Einstein pursued in formulating this theory.

\(^68\) That the theory of relativity is in many aspects the culmination of the programme of classical physics is argued by e.g. Hesse (1961), p. 226.
Einstein challenged Newtonian physics in fundamental ways, the equally correct but neglected point is the number of methodological correspondences with earlier physics [exhibited by Einstein's contributions].

Einstein would have been the first to point out the continuity of his work with classical physics: 'With respect to the theory of relativity it is not at all a question of a revolutionary act, but of a natural development of a line which can be pursued through centuries.'

On such arguments is established the judgement of the theory of special relativity as falling short of constituting a revolutionary innovation, as maintaining continuity with the physical science of the previous century. The theory of special relativity resembles Copernican astronomy in having been designed to rid of imperfections the aesthetic structure of the science forged by the paradigm then dominant in its discipline, and by this means to contribute to that paradigm's fullest development.

5. Aspects of revolution: quantum mechanics

The previous section suggested that the theory of special relativity gained adherents on the strength of its aesthetic features as well as of its logico-empirical virtues. The reception obtained by quantum mechanics was different: the extra-empirical features of the theory proved repellent rather than attractive to the community in which it was formulated, and the endorsement which it acquired was won in virtue of its logico-empirical strengths alone and despite the aesthetic features which it exhibited. This section will allude briefly to the reactions of Planck and Einstein towards quantum mechanics as an illustration of the reception accorded by a community to a theory which constitutes a revolutionary innovation.

The early history of quantum mechanics may be subdivided into three phases. The first opened with the formulation of the quantum theory of black-body radiation in 1900; the second with the quantum atomic theory of Bohr in 1913. Despite their empirical success, it became increasingly apparent in the early

---

69 Holton (1973), p. 170; emphasis in the original.
1920s that the forms of quantum mechanics developed up to that time were
inadequate. They amounted less to a unified and consistent body of theory than
to a piecemeal doctoring of classical physics, a set of conventions governing when
to impose upon classical principles ‘quantum conditions’ restricting the range of
values which given variables could assume. The more uniform and consistent
version of quantum physics was constructed in the third phase through the work
of Heisenberg, Schrödinger and others from 1926 onwards.

Whereas the forms of quantum physics obtained during the first two
phases were theories insufficiently fully-formed to exhibit determinate extra-
empirical or aesthetic characteristics, the new quantum physics possessed
distinctive aesthetic qualities in the sense specified by the present treatment. The
most conspicuous aesthetic feature which the new theory exhibited was causal
indeterminism. It was perception of this quality which provoked Planck’s and
Einstein’s opposition to quantum mechanics.

Both these scientists had contributed greatly to development of the first
form of quantum physics. Planck had inaugurated the new approach by
introducing the notion of the quantum of action to explain the energy spectrum
of black-body radiation; Einstein had demonstrated the power of the quantum
hypothesis to explain the experimental findings of the photoelectric effect. Both
scientists lent their assent to early versions of quantum physics on the strength of
the empirical success of their own contributions and those of others. But for both
Planck and Einstein the unpalatable extra-empirical features of the new quantum
mechanics easily outweighed the empirical success of the theory and ensured
their rejection of it.71

It is not that Planck or Einstein denied the predictive attainments of
quantum physics, and rejected it consequently on empirical grounds: both
repeatedly paid tribute to the predictive success of the new theory. But they drew
attention also to other features of the theory — chiefly its indeterministic treatment
of physical events — which they found unacceptable. There exists compelling
textual evidence for the claim that both Planck and Einstein adopted attitudes of
this description towards quantum physics.

In his Nobel Prize address, where he seeks to retrace the origin and
development of quantum theory, Planck explicitly and at length commends the

71 Planck recalls his opposition to quantum mechanics in his (1948), pp. 43-5. For
background information on Einstein’s resistance to quantum physics see Stachel
(1986), and on that of both Planck and Einstein see J. Bernstein (1973), pp. 153-65.
theory for its empirical success in many and diverse areas of physics. A passage immediately following this review, although closing with the acknowledgement that the ultimate fate of the theory is to be decided on the strength of its empirical performance, lets slip Planck's displeasure:

The difficulties which the introduction of the quantum of action into the well-established classical theory has encountered from the outset [...] have gradually increased rather than diminished; and although research in its forward march has in the meantime passed over some of them, the remaining gaps in the theory are the more distressing to the conscientious theoretical physicist. [...] But numbers decide, and in consequence the tables have been turned.

The fact that Planck should admit that 'the numbers' have won to quantum physics the community's support, and that he should nonetheless voice displeasure at the theory, suggests that he consciously harboured extra-empirical reservations about a theory which he recognized to have demonstrated empirical success. The distinction drawn in Planck's statement between two spheres in which support may be extended or withheld from a theory not only helps to entrench the claim made in the present treatment that scientific practice commonly passes separate logico-empirical and aesthetic judgements about theories, but also illustrates the tenacity of some scientists in mustering opposition on aesthetic grounds to a theory even when it has evidently won the day on logico-empirical grounds.

An article of Einstein's reiterates both Planck's commendation of quantum physics on empirical grounds and his over-riding reservations against it. The rhetorical structure of Einstein's article closely followed that of Planck's lecture: his praise for the wealth of empirical success demonstrated by quantum physics is followed abruptly by criticism of it on quite different grounds:

Experiments on interference made with particle rays have given a brilliant proof that the wave character of phenomena of motion as assumed by the theory does, really, correspond to the facts. In addition to this, the theory succeeded, easily, in demonstrating the statistical laws of the transition of a system from one quantum condition to another under the action of external forces, which, from the standpoint of classical mechanics, appears as a miracle. [...] Even an understanding of the laws of radioactive decomposition, at least in their broad lines, was provided by the theory. Probably never before has a theory been evolved which has given a key to the interpretation and calculation of such a heterogeneous group of phenomena as has the quantum theory. In

---

72 Planck (1922), pp. 13-7.
73 Ibid., p. 18.
spite of this, however, I believe that the theory is apt to beguile us into error on our search for a uniform basis for physics, because, in my belief, it is an incomplete representation of real things [...]. The incompleteness of the representation is the outcome of the statistical nature [...] of the laws.\textsuperscript{74}

Einstein's drawing the distinction between the empirical grounds which warranted an initial favourable appraisal of quantum physics and other grounds on which his overall unfavourable verdict was motivated contributes to establishing the reasons for his opposition as extra-empirical or – in the present treatment – aesthetic. His biographers concur that Einstein's misgivings at indeterminism were based essentially on aesthetic considerations: for him the harmony of the theory would be marred if, to use his own metaphor, God were depicted in it as casting dice.\textsuperscript{75}

The aesthetic opposition raised to quantum mechanics, and the fact that the theory entrenched itself in the community of physical scientists on the strength of its predictive attainments and despite its aesthetic features, suffice on the present account to consider this theory revolutionary. The conduct of members of the community illustrates the behaviour of the factions which on the present account is to be expected in time of revolution. In chapter 5 it was suggested that members of the conservative faction would remain wedded to the aesthetic canons of theory-evaluation bequeathed by the paradigm of which the period of domination is ending, and resist the adoption of the new and aesthetically-innovative theory on the grounds of its violation of those canons. Members of the progressive faction would on the contrary not allow a commitment to past aesthetic canons to dissuade them from embracing the new theory and attaining the greater predictive power which this brings, even if this necessitated suspending the process of theory-assessment on aesthetic grounds. If Einstein is a representative member of the conservative faction which opposed quantum mechanics from 1927 onwards, his counterpart in the progressive faction is Bohr: in their celebrated series of discussions Einstein customarily attacked quantum mechanics on extra-empirical grounds, while Bohr – declining to attempt a point-by-point rebuttal of Einstein's aesthetic misgivings – rested his case on the great and increasing empirical success of quantum mechanics.\textsuperscript{76} In time one should expect the metainductive construction of aesthetic canons of theory-appraisal to

\textsuperscript{74} Einstein (1936), p. 374; emphasis in the original.

\textsuperscript{75} See e.g. the biography of Hoffman (1972), p. 195.

\textsuperscript{76} Bohr's recollections of his discussions with Einstein on quantum mechanics are contained in his (1949).
yield a canon for which a theory’s exhibiting indeterminism no longer counted against the adoption of that theory. This process has perhaps by now been completed: sixty years after the formulation of the new quantum mechanics few physicists continue to offer resistance to its adoption on the grounds that it is indeterministic.
Chapter Eleven
THE DEFINITION OF SCIENTIFIC RATIONALITY

1. Review of the argument

The aim of this treatment has been to contribute some maintenance-work to that hierarchy of partial models of science which is topped by the rationalist image. The problem for the rationalist image of science which was identified at the outset of this work was constituted by its apparent incompatibility with two lower-level members of the hierarchy of currently popular partial models. The first of these lower-level models asserted that science's past development has been fractured by revolutions into periods adhering to norms of theory-formulation and evaluation which are peculiar to each period and not altogether shared by adjacent periods; the second asserted that many important acts of theory-evaluation and choice have been the effect of scientists' applying evaluative criteria which paid regard to the aesthetic as opposed to the logico-empirical features of theories.

The prospect was anticipated that those who with the aid of such hierarchies of models seek to understand the methodology and the history of science might abandon the rationalist image on the grounds of its incompatibility with attractive lower-level models like the two mentioned here, and choose a different model as their image of science of highest generality. In support of the rationalist image, the present treatment aimed to show that the adoption of at least particular formulations of these two lower-level models would not compel the abandonment of rationalist conceptions of science.

The first step in the achievement of this aim was the construction of the particular formulations of the two lower-level models which would ensure this degree of consistency. The fact that the currently available models of scientific revolution are considerably more sophisticated than are current models of the practice of theory-evaluation on aesthetic grounds meant that the greater part of the foregoing treatment had to be directed towards the construction of a model of
the latter phenomenon of scientific practice. As it then happened, the content of
the model of scientific revolutions put forth in chapter 8 was very largely
prompted by the work done earlier in the treatment on the model of aesthetic
time-evaluation.

Both the lower-level models constructed here postulate the existence of a
canon for the evaluation of scientific theories which is composed of two sets of
criteria, distinguished from one another by the origin and the manner of the
variability of the criteria by which they are composed. The criteria of one set have
a metarationalistic origin: they are criteria of which the application in theory-evaluation
and choice is suggested by practical reason to be conducive to the
attainment of the goal which the community imposes on science. Scientific
communities typically characterize this goal in terms of increasing verisimilitude
or observational success, or at any rate in logico-empirical terms. The choice of
logico-empirical terms as those in which to couch a specification of the goal of
science ensures that the methodological precepts which practical reason prescribes
in the furtherance of its attainment will relate principally to logico-empirical
aspects of theories; in particular, if these precepts are formulated as criteria of
theory-choice, the features of theories to which the criteria relate will be their
logico-empirical features.

The manner of the origin of criteria of this set bestows on them a high
degree of stability: barring mistakes in the application of practical reason to the
case, the canon of logico-empirical precepts will not require revision for at least as
long as the community leaves unaltered the formulation which it attributes to the
goal of science. Since changes in the formulation of the goal of science do not
appear to have been frequent in history, the canon of logico-empirical evaluative
criteria are possessed of considerable historical stability.

Two interlocking reasons were adduced for the belief that it was necessary
to supplement reference to this canon of methodological norms by allusion to a
set of quite different criteria of theory-evaluation. The first was the observation
that it was impossible adequately to reconstruct salient episodes in the history of
science by reference to logico-empirical evaluative criteria alone, and hence that a
further set of criteria were required for the completion of the historian’s task. The
second set of considerations, drawn from philosophical aesthetics, suggested that
scientific theories were susceptible to a mode of perception that could be
considered aesthetic in having in view no utilitarian ends, instead turning upon
its object a disinterested gaze. This supposition gained support from the
observation that the procedures historically followed in scientific practice appear
to have included the perception of the aesthetic qualities of theories and their evaluation on nonutilitarian grounds alongside the perception of their logico-empirical features and their evaluation on utilitarian grounds.

The set of aesthetic criteria of theory-assessment of which the operation was postulated by these considerations was held in chapter 7 to possess a metainductive origin, and in consequence to exhibit a degree of historical stability much lower than that of logico-empirical evaluative criteria. The varying degree of accord between the recommendations in cases of theory-choice of the two sets of criteria enabled two states in history of science to be envisaged: a state of normal science, obtained when the recommendations of the two sets maintain a reasonably close accord, and a state of pre-revolutionary crisis, when the joint application of the two sets of criteria renders theory-choice undecidable. The interrelations of the two sets of criteria enabled a mechanism to be proposed in chapter 8 for the advent and termination of scientific revolutions.

2. The scientist's look

From within the present work, scientific practice emerges as the joint application of two modes of perception and of two corresponding classes of evaluative criteria. The first mode of perception, which has hitherto received the lion's share of philosophers' attention, may be termed utilitarian, interested, or transitive. It is utilitarian or interested in the sense that under this mode of attention scientific theories are perceived as vehicles for the attainment of certain ends or goals; it is a mode of transitive perception in the sense that the percipient's attention dwells on the theory in its relationship with entities external to it. Under this mode of perception theories are appraised as the vehicles for the attainment of the goal of science, the eventual formulation of the complete, empirically adequate or true explanatory account of the universe; they are thus perceived not as free-standing constructs but rather in their relationship \textit{vis-à-vis} the phenomena or the data.

The second mode of perception counterbalances the first in these respects. The mode of aesthetic perception is nonutilitarian, disinterested, or intransitive. Under its operation scientific theories are not perceived and appraised as vehicles for the attainment of goals external to them: this mode of perception rather regards theories with an attention which is absolute or detached from
considerations of ends. When a theory is evaluated by a percipient in this mode of attention, it is evaluated for its possession of aesthetic features. To the extent that under this mode scientific theories are perceived as self-sufficient constructs and not in their relationship with the data, this mode of perception is intransitive.

The present treatment thus attributes to the scientist recourse to both the mode of utilitarian perception which is informed by criteria of logico-empirical evaluation, and the mode of disinterested perception which is informed by criteria of aesthetic assessment. The scientific look, the gaze which scientists turn upon the products of their activity, is composed of both these elements.

The suggestion that scientists apply to theories both modes of perception gains support from various sources. It finds backing in historiography, to the extent to which reconstructions of instances of theory-choice which refer to scientists' perception of their theories as both vehicles of utility to be assessed with an end in view and intellectual constructs to be judged on disinterested or aesthetic criteria exhibit a degree of historiographic fidelity greater than that of reconstructions which refuse to acknowledge scientists' application of a mode of attention other than the utilitarian and of categories of evaluation other than logico-empirical. It gains support too from the methodological pronouncements of scientists, who explicitly regard themselves as turning upon scientific theories both an interested gaze which hopes to discern in them empirical virtues, and a disinterested look which searches for purely perceptual or aesthetic quality.

3. The adherence to rationalist precepts

The present treatment undertook to demonstrate the consistency of the rationalist image of science with the two lower-level models which had at the outset appeared to conflict with the broad claims of the top-level model. Chapter 8 showed how both the occurrence of revolutionary discontinuities in scientific progress and the practice of theory-evaluation on aesthetic grounds could be reconciled with tenets of the rationalist image.

The achievement of mutual consistency between a set of models such as these is however apt to perturb or alter the character of each of them. For instance, the model of scientific revolutions which is embraced here is an adaptation of that of Kuhn, which is among all previously-existing models of
scientific practice the one to which the present treatment most directly reacts. But the effect of bringing the three models into consistency has been to alter not only the character of the two lower-level models but also that of their top-level counterpart. The rationalist image of science which emerges from the present treatment possesses features not shared by its formulations at the hands of previous authors.

Scientific rationality finds its bedrock in an application of practical reason, an inference from ends to means. Science is, and is universally perceived by scientific communities to be, a goal-directed enterprise. Rationalist accounts of science have tended to posit that its goal is the formulation of empirically adequate accounts of natural phenomena. A central component of science's methodological canon is drawn up by considering what policies will best achieve this aim. The application of practical reason suggests that the aim will best be attained by acting in accordance with a set of criteria which prescribe that, under conditions of equality, those theories should be preferred which exhibit certain logico-empirical qualities. In the context created by the prescriptive assumptions of science, to act rationally is to act in accordance with these criteria.

If the members of a scientific community successfully resolved that they would allow their actions to be determined only by their canon of logico-empirical precepts, and there existed no other set of precepts which could encourage departures from the behaviour recommended by the logico-empirical canon, then presumably the sole lapses from rational behaviour in science would be due to misinterpretations or misapplications of the logico-empirical criteria of theory-evaluation. But communities are clearly unable to ensure that no precepts which might conflict with their logico-empirical criteria of theory-choice will take a hand in determining their behaviour. Consequently, departures from rational behaviour as this is defined by the prescriptive assumptions of science are more frequent and also more susceptible to reasoned defence than one might otherwise expect, since they may be prompted and justified by the sets of precepts separate from the logico-empirical criteria of theory-choice to which the behaviour of communities may be responsive.

The present treatment has contended that, even if one leaves aside the multitude of other sets of precepts - of varying degrees of externality - to which scientific communities may be subject, there exists one set of methodological precepts to which acts of theory-evaluation and choice on very many occasions in the history of science adhere. This is of course the canon of aesthetic criteria of theory-evaluation. Here is therefore a codification which may systematically
encourage departures from the adherence to logico-empirical precepts, and thus from rational behaviour as this is defined by the prescriptive assumptions of science.

In periods of so-called normal science, the logico-empirical and the aesthetic canons of theory-evaluation will remain according to the present treatment in close agreement. The choices among theories which the community in that phase of history takes, and which will be decided upon by the joint application of both their canons of evaluative criteria, will not differ greatly from the choices which the community would have taken on the guidance of their logico-empirical criteria alone. Thus in periods of normal science the conduct of a community accords closely with the ideal of the rational behaviour as this would be determined by the imposition of the logico-empirical methodological precepts.

At the approach of a revolutionary crisis, on the other hand, the recommendations for theory-choice put forth by the community’s aesthetic canon will diverge increasingly from those simultaneously advanced by their logico-empirical criteria. Some of the choices among theories which the community takes in this phase will be swayed by the aesthetic canon to which it pays allegiance, and will thus depart from the choices which the community would have taken in the absence of its aesthetic canon and under the guidance of the logico-empirical criteria alone. In such circumstances the community’s acts will be said to deviate from the optimally rational behaviour, as this is defined by the prescriptive assumptions of science and consequently by the application of the logico-empirical criteria of theory-choice.

In times of pre-revolutionary crisis, therefore, aesthetic criteria of theory-assessment amount to an influence which perturbs the community’s practice of theory-choice, causing it to deviate from the rational sequence of choices. According to the present model, of course, the deviation from rationality in scientific practice is not allowed to persist indefinitely: in a pre-revolutionary period those members of the community whose choices are taken on aesthetic grounds are thereby led to embrace theories which are empirically less successful than are the theories adopted by those whose choices are grounded on logico-empirical grounds. As chapter 8 illustrated, the empirical inferiority of the choices taken on aesthetic grounds will grow increasingly manifest until all members of the community agree that the constraints on theory-choice imposed by the aesthetic evaluative canon ought to be relaxed. By this decision, the community’s practice of theory-choice will be returned to full rationality as this is defined by the application of logico-empirical considerations; this state will persist into the
next period of normal science, until the aesthetic canon of theory-assessment which the community constructs again enters into conflict with the recommendations of logico-empirical criteria.

4. The memory of science

The difference between the two canons of criteria of theory-evaluation of which the use by scientists is invoked by the present treatment may be summarized in the following way: whereas logico-empirical criteria are of a priori formulation and express what scientific communities hold to be unvarying or at least very slowly varying requirements of theories, aesthetic criteria are formulated on metainductive considerations. As chapter 7 argued, a community assembles its aesthetic canon at a certain date from among the aesthetic features of all past theories by attributing to each feature a degree of favour proportional to the degree of empirical success scored to that date by the set of theories which have appeared to embody it: the community’s aesthetic canon is then composed of the set of mutually consistent such features which have gained the greatest favour.

Clearly, a crucial step in the construction of an aesthetic canon of theory-evaluation is the formulation of a judgement about the past. In order to assemble a canon of aesthetic criteria for theory-assessment, the eye of the scientific community must range over the historical record, sort by their aesthetic features the theories which the community has at some past time embraced, and attribute to each such feature a degree of favour proportional to the perceived empirical success of the set of theories which have exhibited that feature. Scientific communities thus have to discharge the historiographic task of isolating and evaluating the theories which they have in the past embraced.

As the present treatment has made plain, the application of the canon of aesthetic criteria which is constructed by these means is a important component of the community’s scientific practice. The fact that this canon is formulated partly on the basis of historiographic judgements entails that the norms of scientific practice are constructed in part on historiographic grounds. In other words, scientists appear to derive their practice partly from a view of the history of science and a notion of which instances in that history have been methodologically most notable. Scientific judgement thus comes to be determined
in part by historical judgement: historiography may be seen as a contributor to
the construction of scientific methodology. The scientific community retains a
historiographic ‘memory’ of instances of theory-choice and on their basis
constructs the canon of practice which is to guide it further.

This view attributes to the scientist important historiographic interests. It
predicts that practising scientists will find the history of their discipline a
necessary subject of study. Their stake in the study will differ from that of the
professional historian. While the aim of the historian in retracing science’s past
will be to understand the development of science or perhaps of wider culture for
its own sake, the aim underlying the scientist’s approach to the history of his or
her discipline will be to acquire the data necessary to the construction of the
inductive component of his or her methodological canon.

Naturally much of the historiographic investigations and judgements of
practising scientists will remain unwritten or implicit: after all, they are judged by
the eventual outcome in science of their contributions and not, as are the
historians, by their strictly historiographic writings. Nonetheless, there is much
evidence that scientists turn to history for guidance on precepts of method or for
didactically valuable instances of scientific judgement. Such evidence tends to
support the contention of the present treatment that data from the history of
science plays an important role in the construction of scientific method, and that
scientists themselves show both awareness of this use and alacrity in exploiting it.

Examples of the scientist turning historiographer in the expectation that
history will offer methodological guidance in his or her present researches
abound. Newton and many eighteenth-century scientists thought it appropriate to
preface scientific works with historical notes locating their own research in a story
of natural-philosophical inquiry stretching back many years.

This view attaches to the history of science an importance much greater
than even that which it receives from some of its most dedicated recent partisans.
Kuhn stated that the history of science could serve a role in the understanding of

---

1 The reluctance of the practising scientist to commit to print his or her
historiographic judgements has moreover naturally increased with the creation and
increasing formalization of a professional discipline of history of science.

2 For evidence of Newton’s historiographic interests see his *Scholia classica* the text of
which is reproduced with commentary in Casini (1984); for further study of
Newton’s general interest in historiography see Manuel (1963). Galluzzi (1988) has
charted the extensive historical comments into which French scientists of the
eighteenth century set their new contributions. In general on the role of scientists’
historiographic remarks in their construction of scientific methodology see Kragh
scientific methodology.\textsuperscript{3} From the findings presented here, it emerges that the history of science contributes essentially not only to the understanding of scientific methodology but to its very construction.

\textsuperscript{3} Kuhn (1962), p. 1.
BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


Johnson, R. 1982. ‘Parsimony principles in phylogenetic systematics: a critical re-
appraisal', in *Evolutionary Theory* 6, pp. 79-90.


BIBLIOGRAPHY

Cambridge, Mass.: Harvard University Press.


BIBLIOGRAPHY

Cambridge, Mass.: M.I.T. Press.


208
BIBLIOGRAPHY

of Karl Popper. La Salle, Ill.: Open Court; pp. 3-181.


Semple, R.B. 1976. ‘Tate Gallery buys pile of bricks – or is it art?’, in The New

Shaftesbury, A.A. Cooper, Third Earl of, 1711. Characteristics of Men, Manners, Opinions, Times etc. Edited by J.M. Robertson. Two volumes. London: Grant Richards, 1900.


BIBLIOGRAPHY


