255 BOSTON STUDIES IN THE PHILOSOPHY OF SCIENCE

Rethinking Scientific Change and Theory Comparison:

Stabilities, Ruptures, Incommensurabilities?

Edited by
Léna Soler, Howard Sankey and Paul Hoyningen-Huene

Springer

RETHINKING SCIENTIFIC CHANGE AND THEORY COMPARISON

BOSTON STUDIES IN THE PHILOSOPHY OF SCIENCE

VOLUME 255

RETHINKING SCIENTIFIC CHANGE AND THEORY COMPARISON: STABILITIES, RUPTURES, INCOMMENSURABILITIES?

Edited by

LÉNA SOLER
LPHS-Archives Henri Poincaré, Nancy, France

HOWARD SANKEY
School of Philosophy, University of Melbourne, Australia

and

PAUL HOYNINGEN-HUENE
Center for Philosophy and Ethics of Science, University of Hannover, Germany

Springer

Library of Congress Control Number. 2008920285

ISBN 978-1-4020-6274-2 (HB)
ISBN 978-1-4020-6279-7 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

CONTENTS

CONTRIBUTORS

Aristides Baltas
National Technical University, Athens, Greece

Anouk Barberousse
Institut d'Histoire et de Philosophie des Sciences et des Techniques, Paris, France

Soazig Le Bihan
LPHS – Archives Henri Poincaré, Nancy, France and Illinois Institute of Technology, Chicago, IL, USA

Alexander Bird
University of Bristol, United Kingdom

Michel Bitbol
CREA, CNRS/Ecole Polytechnique, Paris, France

Martin Carrier
Department of Philosophy, Bielefeld University, Germany

Steve Clarke
Program on the Ethics of the New Biosciences, James Martin 21st Century School, University of Oxford, United Kingdom, and Centre for Applied Philosophy and Public Ethics, Charles Sturt University, New South Wales, Australia

Bernard d'Espagnat
Institut de France, Paris, France

Ronald N. Giere
Center for Philosophy of Science, University of Minnesota, Minneapolis, MN, USA

Rom Harré
Linacre College, Oxford, United Kingdom and Georgetown University, Washington, DC, USA

Stephan Hartmann
Center for Logic and Philosophy of Science, Tilburg University, Tilburg, The Netherlands

Paul Hoyningen-Huene
Center for Philosophy and Ethics of Science, University of Hannover, Germany

Edward Jurkowitz
Department of Humanities, Illinois Institute of Technology, Chicago, IL, USA

Igor Ly
LPHS – Archives Poincaré, Université Nancy 2, CNRS, Nancy, France

Thomas Nickles
University of Nevada, Reno, USA

Robert Nola
Department of Philosophy, The University of Auckland, New Zealand

Eric Oberheim
Humboldt University, Berlin, Germany

Howard Sankey
School of Philosophy, University of Melbourne, Australia

Léna Soler
LPHS – Archives Henri Poincaré, Nancy, France

Mauricio Suárez
Department of Logic and Philosophy of Science, Complutense University, Madrid, Spain

Paul Teller
Department of Philosophy, University of California, Davis, CA

Emiliano Trizio
LPHS – Archives Henri Poincaré, Nancy, France

Marcel Weber
University of Basel, Basel, Switzerland

Hervé Zwirn
Institut d'Histoire de Philosophie des Sciences et des Techniques, and Ecole Normale Supérieure de Cachan, Paris, France

INTRODUCTION

LÉNA SOLER

This volume is a collection of essays devoted to the analysis of scientific change and stability. It represents the most recent thinking on the topic of incommensurability and scientific theory change. It explores the balance and tension that exists between commensurability and continuity (or stabilities) on the one hand, and incommensurability and discontinuity (or ruptures) on the other. And it discusses some central epistemological consequences regarding the nature of scientific progress, rationality and realism. With respect to these topics, it investigates a number of new avenues and revisits some familiar issues, with a focus on the history and philosophy of physics, in a way that is informed by developments in cognitive sciences as well as the claims of "New experimentalists".

The essays in this book are fully revised versions of papers which were originally presented at the international colloquium, "Repenser l'évaluation comparative des théories scientifiques: stabilités, ruptures, incommensurabilités?" organized by Léna Soler and Paul Hoyningen-Huene at the University of Nancy, France, in June 2004. Each paper is followed by a critical comment, which either represents an opposing viewpoint or suggests some developments. The conference was a striking example of the sort of genuine dialogue that can take place between philosophers of science, historians of science and scientists who come from different traditions and endorse opposing commitments. I hope that this is evident in the book too and that it will constitute one of its attractions. The book is also unique in reflecting and promoting interactions between French philosophy of science and Anglo-American philosophy of science.

As an introduction, I will describe the way the problem of scientific change has been framed and transformed in the philosophy of science throughout the twentieth century and up to the present, emphasising general tendencies in the way problems have shifted, and indicating how the different contributions of this book are related to each of these issues.

The twentieth century has been the theatre of important scientific transformations – so important that they have often been described as ruptures, revolutions or mutations. These transformations manifested themselves at different levels: at the level of high-level theories; of scientific instrumentation; of experimental practices; of the organisation of scientific research. Philosophers of science have sought to characterize these changes, to understand the reasons for them and to explore their implications.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 1–17.

1. SCIENTIFIC CHANGE AS RUPTURE BETWEEN THEORIES

What struck philosophers of science first, and what has been at the centre of the debates during the latter two thirds of the twentieth century, is *rupture at the level of theories*. Indisputably, a number of new scientific *theories* with unexpected characteristics emerged during this period, which broke ontologically and methodologically with earlier theories. This was most notably the case in physics with the theories of relativity and, especially, quantum physics. The result was a considerable enrichment of the range of theories and inter-theoretic relations available for examination. And the comparative evaluation of competing scientific theories, understood as one of the crucial tasks of working theoretical scientists, became an increasingly important problem for philosophers of science.

From an epistemological point of view, what was at stake, in that context, was nothing less than scientific realism, scientific progress, scientific rationality and relativism. A number of scholars, reflecting on these perceived-as-deep transformations of scientific contents and criteria of theory-choice, came to conclude that traditional scientific realism was untenable. They came to endorse weaker and weaker conceptions of scientific progress (so weak that their foes equate them with the claim "*no* scientific progress"), as well as to deny that there may be rational grounds for the judgement that one theory is objectively better than another.

Such alleged dramatic consequences of radical ruptures arising in the theoretical sphere have usually been discussed under the heading of "incommensurability". Two kinds of incommensurability have been progressively recognized as different in principle, although most of the time intertwined in real historical cases. They are commonly labelled "semantic incommensurability" and "methodological incommensurability".[1]

2. THE PROBLEM OF SEMANTIC INCOMMENSURABILITY

Semantic incommensurability involves radical changes at the level of theoretical contents that manifest an incompatibility not reducible to a simple logical contradiction. The incompatibility is apparently related to differences in the semantic resources themselves, so that one is inclined to say that there is no common measure at the very level of what is thinkable or expressible in symbols. Such incompatibility has been characterized by the late Thomas Kuhn, and is currently described in recent works, as the impossibility of translating some key words of one theory into the vocabulary of the other, or as a non-homology between the taxonomic structures of successive paradigms.

At the level of the incommensurability of theoretical contents, the task is to achieve a fine-grained characterization of the differences involved, as well as to discuss how

[1] For recent overviews on the incommensurability problem, see Soler (2000, chap. VII, 2004), Hoyningen-Huene and Sankey (2001).

deep the differences encountered in the actual history of science are, and why these differences have opened up.

The most obvious and most debated epistemological issue of semantic incommensurability is the problem of scientific realism. The simple claim that there are scientific revolutions at the descriptive level, i.e., the claim that there are, in the actual history of science, important ruptures at the level of "what science tells about the world", is, in itself, already a threat to any form of "correspondentist" or convergent realism (i.e., the thesis that the content of successive scientific theories approximately corresponds to the world and converges ever more toward the truth). Yet, the more precise claim that these descriptive scientific disruptions originate in deep differences rooted in the very resources of what is expressible, is far more subversive, since it throws into question the very formulation of the problem of the relation between theories devised by humans and their non-human referent. At least, it forces us to face seriously the idea of a constitutive power of language in science.

3. THE PROBLEM OF METHODOLOGICAL INCOMMENSURABILITY

Methodological incommensurability involves an irreducible incompatibility between the norms that underlie alternative programs of scientific investigation or traditions of research, for example, between the (often tacit) commitments about what is a genuine scientific problem, solution or proof. The most extreme case would be the situation in which no standards and values are shared between two traditions, each considering itself as scientific.

At the methodological level, the task is (1) to identify the efficient standards involved, which are, in normal conditions, largely tacit, intuitive in application, and by nature not precise; (2) to understand the way the different standards are or may be related to each other and to semantic differences between theories; (3) to evaluate how deep are or may be the methodological transformations encountered in the real practices, and to discuss why these transformations have occurred. At a more general meta-methodological level, the task is also (4) to reflect on the aim, the fruitfulness and the limits of any methodological theories of science.

The central epistemological issue commonly associated with this methodological incommensurability is the problem of scientific rationality and relativism. Relativism is here understood as the problem of knowing whether scientists have at their disposal, in each stage of scientific development, sufficiently good uniformly compelling reasons for deciding what is better at the level of validation procedures.

4. SCIENTIFIC CHANGE ENLARGED AND REFRAMED IN TERMS OF RUPTURES BETWEEN SCIENTIFIC PRACTICES

This is the traditional way to frame the problem of scientific change. The debate about these issues continues to be lively today. However, important shifts have taken place at the level of the interests of philosophers of science as well as historians, sociologists and ethnologists of science, and, correlatively, of the problem formulations. While interests and formulations were first mainly, if not exclusively, centred *on full-blown theories* usually understood as sets of explicit statements, more and more other dimensions have been claimed to be essential to the understanding of science and have been integrated into the picture as the investigation has continued. These include non-linguistic and partly tacit aspects such as know-how; experimental actions and material aspects constituting laboratory life; more or less local standards and values; commitments and projects of all sorts; not to mention psychological or social factors. Following what has been called the "turn to practice", the original issues have been re-framed in terms of scientific (theoretical or experimental) practices, if not in terms of cultural traditions socially labelled "scientific".

As a result, theories have progressively ceased to be reduced to sets of explicit statements, and scientific change has ceased to be thought of only in terms of theory change. The "New Experimentalists", in particular, have denounced the exclusive focus of philosophers of science on theories, representations, truth appraisal and epistemic issues. Particularly representative of these critics and their rhetoric is the contribution of Thomas Nickles in this volume. Paul Teller's paper, which adopts the "modelling approach to science", is also a case in point. It argues that most of the traditional issues related to incommensurability vanish if the philosopher stops taking truth as the aim of science and is instead attentive to a multiplicity of local and often pragmatic goals and values. As Ronald Giere points out in his commentary, Teller's conception leads us to be especially attentive to practical virtues in methodological appraisal – not just to cognitive values, or worse not just to one of them, the over-valued "truth" or "proximity to the truth". Real practitioners often evaluate the comparative merits of different available models on practical grounds (e.g., computational tractability). Anouk Barberousse, whose approach is more akin to the analytic tradition, also joins those who regret that philosophers of science have focused so much on theories as products. She wants to take into account the "first person use of theories". In this perspective, scientific theories are not just viewed as sets of objective scientific sentences that one can find in textbooks. They are also considered as sets of individually interpreted sentences.

As a special case of this general tendency, the linguistic characterization of incommensurability has been criticized as too narrow, dispensable or even misleading. In this volume these criticisms are articulated in Nickles' and Teller's articles. They are also echoed in several other papers. Bird, for example, stresses that "world change" is richer than a change at the level of meaning, that the focus on theories, concepts and language has been misplaced, and that although one side of incommensurability concerns understanding, incommensurability can perfectly well be characterized without recourse to meaning or linguistic issues. Aristides Baltas offers another illustration. By contrast with Bird, he finds the characterization of incommensurability in terms of concepts and language especially relevant (although he explicitly recognises that scientific practices are never merely linguistic). But the linguistic characterization of incommensurability he proposes is grounded in a shifted, enlarged conception of language inspired by Wittgenstein's famous dictum that "meaning is use". In such a framework, language is understood as an activity and the performative power of words is integrated into the picture.

Correlatively, the methodological side of incommensurability has been broadened, in the sense that conflicts between incompatible standards and values have been exhibited at many other levels than at the level of theory appraisal – although such cases have not always been discussed under the heading of methodological incommensurability. Several New Experimentalists, in particular, claim that there are significant conflicts of this kind at the experimental level, notably regarding what different practitioners consider to be a convincing experimental demonstration or a robust piece of evidence (e.g., Pickering, 1984, 1995; Galison, 1997). My contribution in this volume discusses aspects of these issues, and suggests that such cases should be recognized as local instances of methodological incommensurability (or "value incommensurability" to use the expression favoured by Giere in this volume).

Many authors urge, in the spirit of Kuhn's early writings, the need to examine the nature of scientific education in order to understand science, because specific features of this education appear responsible for essential characteristics of science. In this vein, scholars have especially insisted on learning processes which involve repetitive practice with exemplars. It is by such processes, they argue, that the scientific apprentice acquires the efficient (partly tacit) know-how and the background knowledge that render him able to solve new scientific problems by analogy with previously incorporated solutions. This has led to paying special attention to the tacit dimension of scientific building and scientific judgments.

Such a tendency is exemplified in several contributions of the book. For example, Bird suggests that incommensurability may be understood as differences in largely unconscious "quasi-intuitive cognitive capacities" which involve tacit background knowledge and commitments. He also relates the nature and efficiency of these quasi-intuitive cognitive habits to the process through which they have been learned, namely, repetitive training and exposure to standard exemplars. Baltas also makes room for unconscious background elements and is interested in educational aspects. In his Wittgenstein-inspired characterization of semantic incommensurability, he puts forward the hidden assumptions that constitute linguistic uses in general and scientific practices in particular. These background assumptions are tacit, self-evident and unquestioned ingredients that play the role of Wittgensteinian hinges and silently drive the understanding and the correct use of concepts. This is true for professional scientists as well as for students. Baltas attempts to draw the pedagogical implications of such a situation. He examines the resources that scientific professors have at their disposal when they teach new incommensurable concepts and theories to students who are entrenched in another incommensurable scientific paradigm. His proposal is to use nonsense, that is, to build bridge-sentences that are literally nonsensical but that do something (that are "performatively meaningful"). In his comment on Baltas' paper, Eric Oberheim enriches Baltas's proposal with other pedagogical resources that also involve pragmatic and tacit aspects, such as the use of analogies and metaphors. Barberousse's first person perspective also gives an important place to implicit aspects. In this conception, the different versions of the same scientific theory are analyzed as sets of sentences interpreted by different individual scientists from their own singular perspective, and the process of this individual appropriation – notably the process of individual understanding and development of concepts – is described, drawing on Robert Brandom's work, as a special kind of activity, namely, the activity of "making a claim explicit". The elaboration of a new version by a particular scientist is described as the act of making explicit, in an original manner, a theoretical content that, in a sense, was already there, although implicitly. This act, Barberousse insists, is not at all obvious or "automatic". It requires talent, creativity and a lot of know-how.

This being said, the emphasis on science as an activity, and the new interests that go with it, do not eliminate or dissolve traditional issues previously associated with linguistically and theoretically centred approaches. In particular, the central problem of theory comparison has not disappeared. This is why "theory comparison" is mentioned in the title of the book. It remains a genuine and important philosophical task, even if it does not exhaust the problem of scientific change and if its characterization must take into account other aspects of scientific practices than theories considered in isolation.

5. STILL ALIVE TRADITIONAL ISSUES RELATED TO THEORY COMPARISON AFTER THE TURN TO PRACTICE

The traditional problem of theory comparison, as it developed following the introduction of incommensurability by Kuhn and Feyerabend in the 1960s, has two related faces. Semantic incommensurability challenged the very possibility of a point-by-point comparison of subparts of old and new incommensurable theories, and called for a precise characterization of the relationships between alternative or competing theories. Methodological incommensurability challenged the universal, compelling and objective status of the criteria according to which practitioners evaluate the relative merits of competing theories. These two interrelated faces of theory comparison are still active problems, and many contributions to this volume are related to them.

With respect to the kind of relations that exist between successive theories, Stephan Hartmann argues that significant continuities always hold, although only partial and often complex ones, even when "scientific revolutions" occur. He elaborates a repertory of different types of partial correspondences between theories. This leads him to revisit the difficulties and ambiguities of the correspondence relations.

Robert Nola argues that even if a point-by-point comparison between semantically incommensurable theories is impossible, this is not at all a problem as long as we can separate the incommensurable theories from an independent set of reproducible observational effects consensually recognized as 'the effects these different theories are about'. In such cases, frequent in the history of science according to Nola, the stable set of reproducible effects points to the same thing (or possibly set of things) as their cause (despite the fact that the incommensurable theories conceive the thing or things very differently). In this manner, referential continuity is preserved despite drastic theoretical changes.

As for Barberousse, she insists that interest in scientific revolutions has led to a neglect of non-revolutionary developments, with the result that we are not well-equipped to think about the relation between different versions of one and the same theory. In other words, she addresses the issue of theory comparison in the case of intra-theoretical change (comparison between theoretical versions of a unique theory). Her problem is related to Kuhn's "essential tension": the aim is to reconcile the idea that different versions are essentially similar (otherwise they would not be viewed as versions of the same theory) with the intuition that a version can nevertheless introduce significant novelty and require great creativity. It is in order to make progress with this problem that she appeals to Brandom's notion of "making a claim explicit".

The difference between intra-theoretical and inter-theoretical changes, or between non-revolutionary and revolutionary scientific episodes, reactivates traditional unsolved difficulties that can be grouped under the heading of what I would call "theory individuation". As Igor Ly stresses in his comment on Barberousse's paper, the borderline between cases of the type "from one theory to a different theory" and cases of the type "from one version to another version of the same theory" is not so clear. In this respect the notion of "making a claim explicit" does not help as long as no indication is given about its conditions of application.

Problems related to theory individuation are also highlighted by Edward Jurkowitz in his comment on Hartmann's paper. The questions: "what a theory consists in?", "What a theory 'says'?", and "How far a theory must deviate from a 'core' in order to be a new (possibly incommensurable) theory or even a distinct theory?", are indeed, Jurkowitz argues, very difficult questions that must be decided before considering Hartmann's question "what are the correspondence relations between two theories?". In any case, in order to be in a position to answer, one has to specify when and for whom the answer is supposed to hold. This answer will in general differ for different periods and different subjects.

Several contributions to the book revisit the other face of theory comparison, that is, the comparative evaluation of competing theories and the power of criteria of theory choice. As has been well-known at least since Duhem, empirical and logical criteria are not sufficient to impose unambiguously a unique choice between a set of competing theories. Rom Harré attempts to fill the gap by a criterion of ontological plausibility. He argues that decisions of real practitioners, more especially decisions about which theory is the best candidate with respect to a realist reading, are indeed determined by such considerations, that is, by a requirement of ontological continuity. The point is advocated in the framework of an understanding of scientific activity as an activity of building models, with a special accent on iconic models. In this framework living scientists elaborate models about unknown entities by analogy with known observable entities, and this activity, as well as comparative judgments concerning the merits of different available models, are guided by the maxim: 'prefer the model coordinated with the ontology that is in best harmony with the actually admitted ontology'.

In her comment on Hervé Zwirn's paper, Soazig Le Bihan considers a contemporary, still unresolved case of theory choice which deeply involves the ontological preferences of practitioners: the choice between different (empirically equivalent) scenarios associated with the mathematical formalism of quantum physics. How will realistically inclined physicists be able to choose, for example, between an Everettian scenario and an (incompatible) Bohmian scenario?

With respect to the comparative evaluation of two semantically incommensurable paradigms, two contributors suggest that some objective asymmetries exist, which indeed crucially constitute in fact, and should constitute in principle, judgments of superiority among paradigms. In this vein, Giere argues that the Einsteinian theory is superior to the Newtonian theory, in the sense that on the basis of the former we can understand why the latter works as well as it does in its domain, or why it is impossible to shield gravitational forces, whereas the reverse understanding is not possible. In a similar vein, although within a very different framework, Baltas insists that the new vantage point of the post-revolutionary grammatical space allows a reinterpretation of (some aspects of) the pre-revolutionary one, whereas the reverse is blocked.

Martin Carrier also tackles the problem of underdetermination, but with respect to its implications for philosophical theories of scientific method. He first argues that both Duhemian and Kuhnian underdetermination of theory choice are inescapable and will thus never be solved by any methodology. This leads him to articulate the idea that we should modify the traditional conception of methodology. The aim of a methodological theory should no longer be identified as the exhibition of definite criteria that would unambiguously indicate the best theory among a set of competing ones – this would amount to assigning an unattainable aim to methodologies. Instead, Carrier suggests that the aim of a philosophical theory of scientific method should rather be to order, to classify and to make explicit the mutual relations of the multiple criteria used by scientists.

Faced with the same problem, Paul Teller takes another road. He argues that the problem is artificially generated by a misplaced exclusive valorisation of truth as the aim of science. According to him, traditional issues related to methodological incommensurability vanish once the philosopher renounces the ideal of exact truths, admits that science will never provide anything else than literally false idealizations, and understands that such idealizations give scientists exactly the same practical means as approximate truths would do. In such a perspective, the philosopher becomes convinced that it is desirable to develop a multiplicity of idealizations: indeed, each will have specific virtues, and altogether they will provide a high diversity of resources to the scientific community. In such a perspective the philosopher no longer understands theory choice as a definitive choice in favour of a unique theory with respect to its better proximity to truth, but as a multiplicity of contextual evaluations with regards to the respective capacities of available models as good means for achieving specific (often practical) ends: in Giere's words he becomes a theoretical pluralist. In the framework of such an instrumental rationality, a model that is elected as the best with respect to a given particular aim can be dismissed with respect to other ends. Because contextual instrumental evaluations of this kind are, according to Teller, most of the time consensual and unproblematic, the underdetermination of theory choice by experimental and logical criteria, or the impossibility of justifying criteria in the context of conflict between practitioners, are no longer a problem, or at least they become tractable practical problems.

6. THE SHIFT OF ISSUES ASSOCIATED WITH SCIENTIFIC CHANGE IN THE CONTEXT OF AN EMPHASIS ON LOCAL AND CONTINGENT ASPECTS OF SCIENTIFIC PRACTICES

The net effect of the move toward practice has been, at the level of general philosophical tendencies, an emphasis on local and contingent aspects of science.

The sensitivity to local and contextual features of science is manifested in several papers in the book. Nickles, for example, recommends a pragmatic and flexible conception of scientific method and points to the methodological opportunism of practitioners. Hartmann himself, although in search of a general Bayesian formalization of scientific development, describes the methodologist's task as the construction of a "toolbox" from which scientists may draw when needed, and insists that general methodological principles show their fecundity only when adapted to specific problems. His commentator Edward Jurkowitz, as an historian of science "for whom 'practice' has become a keyword, and who regularly insists on the determinative force of local and even culturally specific forms of theoretical practice", welcomes such declarations.

The issue of contingency is also considered in the book, although to a lesser extent. In the context of her reflection on intra-theoretical change in the case of the history of mechanics, Barberousse asks whether this scientific development was inevitable or not. She attempts to think through the relation between, on the one hand, the objective content of the theory – already there but only implicitly and potentially – which would be responsible for the inevitable aspects of scientific development, and on the other hand a subjective appropriation of this content by individual scientists, which would be responsible for the contingent aspects of scientific development. Following Ian Hacking, my own paper describes the opposition between contingentism and inevitabilism, and identifies it as one of the implications of a possible incommensurability at the experimental level. Other papers encounter this issue in a more peripheral way (see, e.g., Nola Sect. 3.2 or Nickles Sect. 2).[2]

The stress on local and contingent aspects has thrown into question the pretension of science to universality, be it at the level of scientific results or at the level of scientific methods. New varieties of relativist positions have emerged. The possibility of a general picture of science has been questioned in principle. Some have endorsed a radical historicity. In any case, the multiplication of very particular case studies, while it has clearly enriched our approach to science, has correlatively increased the difficulty, for philosophers of science, of the elaboration of a credible global understanding of science. At the same time, many such philosophers think that this is the most important challenge in the current situation.

Many contributors to the volume seek such a global understanding. Bird's account of incommensurability as differences in cognitive habits is indeed a very general picture that purports to be valid for all cognitive processes, included those occurring

[2] On the contingency issue, see L. Soler and H. Sankey eds. (forthcoming). This is a symposium devoted to the contingency issue, with contributions from Allan Franklin, Howard Sankey, Emiliano Trizio and myself.

outside of science. Hartmann's explanation of continuity in science with the help of a Bayesian framework also vindicates a very general validity. Even if the point is questioned by his commentator, Hartmann's research program illustrates the enduring search for general characterizations of science. Teller's and Giere's pluralist and perspectivist picture driven by an instrumental rationality can also be viewed as a description of very broad scope. As for Michel Bitbol, in his comment on Carrier's paper, he endorses a neo-Kantian stance which involves a quasi-universal methodological principle characteristic of the scientific enterprise: the quest for maximal objectivity in a Kantian sense, that is, something closely linked with the search for maximal invariance. The principle takes its quasi-universal status from the vindication that it is built into the very idea of science, that it is constitutive of the very project of science and can thus be seen as one of its conditions of possibility.

Along with the emphasis on contextual and contingent determinants in science, the move toward practice has raised new questions and altered the answers to traditional ones. Here are some examples. What is the exact function of high-level theories, and of competition between them, in scientific research? Are they only one element among many others, no more important than others or indeed, sometimes at least, far less important? How should scientific progress be redefined to take into account the contextual, local and often contingent aspects of scientific practices? When material things, actions, future-oriented concrete projects involved in the practice of science are put forward to the detriment of theoretical representations, are we forced to leave traditional varieties of scientific realism?

7. REALISM VERSUS ANTIREALISM: OLD AND NEW FORMS

Scientific realism remains a central issue in the philosophy of science. Already before the turn to practice, many Kuhnian-inspired philosophers claimed that realism was no longer credible as a consequence of the existence of semantic and methodological incommensurabilities. Within the pragmatic orientation, new and often weaker forms of realism have been introduced, for instance Hacking's and Cartwright's "entity realism", or Pickering's "pragmatical realism" (Hacking, 1983; Cartwright, 1986; Pickering, 1989, 1995). The realist issue, in some of its new forms, and the division between realists and anti-realists, are well represented in the book.

On the side of realist-inclined philosophers, we find notably Rom Harré, Robert Nola, Steve Clarke and Howard Sankey. In his contribution to the present volume, Harré actually does not directly argue for realism. But clearly, a philosopher who accepts the theses he advocates in his paper is in a position to consider the realist stance with sympathy, in agreement with the fact that this is the stance Harré in fact favours. In particular, Harré claims that a realist interpretation of quantum physics, grounded in an ontology of causal powers, is possible. This being said, the theses Harré articulates in his paper do not imply realism, as Suárez insists in his comment. Harré's description of scientists as realists, and his description of ontological plausibility as a criterion that guides model building and the comparative evaluation of models, do not imply, or even strongly favour, the metaphysical thesis of scientific realism. The requirement of ontological continuity, and the ontological stability that the history of science displays as a result, can, Suárez argues, be explained on other grounds in an instrumentalist spirit. For example, one can invoke convenience and economy, that is, reasons that are related to features of human cognition rather than to properties of the world. In other words, the fact that practitioners favour realist readings of models and use criteria of ontological plausibility can be viewed as a good scientific policy (given some characteristics of human cognition) but is not a guarantee that scientific theories indeed pick out components of an independent world.

Robert Nola does not directly address the realist issue either, but he argues for a thesis that is traditionally used to support scientific realism, namely denotational and referential continuity. Taking this for granted, his commentator, Steve Clarke, advocates a particular form of realism, entity realism. According to Nola, the entities that are denoted by theoretical terms can be identified as (roughly) the same, even when radical changes occur in scientific theories. Hence we can conclude that most of our successive past scientific theories refer to the same entities despite the fact that these theories assume radically different conceptions of the entities involved. On this basis, an optimistic meta-induction (OMI) leads us to admit that the same will continue to hold for future theories. Admitting these conclusions, Clarke dissociates OMI and realism about entities on the one hand, and OMI and realism about theories on the other hand, and argues that denotational continuity through theoretical revolutions, at the same time supports entity realism and dismisses realism about theories. Following Hacking and Cartwright, he relates referential success to the possibility of an experimental manipulation of the corresponding entities.

As for Howard Sankey, he claims, in his comment on my paper, that semantic incommensurability, because it is only local, poses no threat to scientific realism. Paul Hoyningen-Huene expresses a different view in his commentary on Bird's paper. In the context of a neo-Kantian framework according to which subject-sided features of scientific theories can never be subtracted from object-sided features, he expresses his conviction that incommensurability, through the experience of world change, essentially challenge and, according to him, has "seriously undermined" any form of scientific realism. Hervé Zwirn, for his part, discusses some aspects of a realist interpretation in the special case of quantum physics equipped with the decoherence theory. His more general position, which is only sketched in the present paper but argued in several writings, is that quantum physics cannot be interpreted as an (even approximate) description of an independent reality. Soazig Le Bihan lists the special difficulties one has to face in order to sustain a realist interpretation of the quantum theory, given, on the one hand, the tension between the uniqueness of measurement results and the multiplicity of the superposed states of the formalism, and, on the other hand, the fact that several empirically equivalent ontological scenarios can be associated with the formalism. Michel Bitbol, who in company with Hoyningen-Huene favours a neo-Kantian orientation, starts with the claim that scientific realism lacks a solid methodological ground, and from there concludes that (approximate) truth must be integrated into our account of science not as an actual achievement but as a regulative ideal.

8. DISCOVERING NEW ASPECTS OF SCIENCE: PRACTICAL REVOLUTIONS? NEW KINDS OF INCOMMENSURABILITY?

The move toward practice has led a number of philosophers of science to claim that it has shed light on important but previously invisible aspects of science. In this vein, Nickles' contribution argues that when we are attentive to pragmatic and future-oriented aspects of science, we are led to discover new kinds of scientific revolutions: disruptive changes that do not obey the Kuhnian scheme. These revolutions are practical disruptions: they are competency destroying. They are ruptures at the level of skills, practical standards, practical expertise and heuristic appraisal. They involve what Nickles calls a "recognition problem": before and after, what appears relevant, fruitful, reliable and legitimate to the field can change drastically. Such disruptions, Nickles claims, often go with a new mapping of the speciality areas and may be relatively independent of theoretical revolutions. In his comment, Emiliano Trizio specifies the characterization by distinguishing two components of Nickles' recognition problem: conflicts between judgments of relevance and conflicts between judgments of fecundity. Only practical disruptions that involve disagreements at the level of relevance might induce the re-structuring of scientific specialties.

In a similar vein, "New Experimentalists" such as Andrew Pickering and Ian Hacking have argued that there is a new form of incommensurability that is found at the level of experimental practices.[3] This form of incommensurability has been overlooked by traditional philosophers of science because of their theory-dominated orientations and exclusive concern with language-based characterizations of incommensurability. Such an incommensurability obtains between scientific practices that have stabilized on the basis of different measurement procedures, instruments, machines and laboratory practices. As a result, the practices are, in Hacking's terms, "literally" incommensurable, in the sense that there is, properly speaking, no shared physical measure between them. (Hacking talks about a "literal incommensurability" and Pickering about a "machinic incommensurability"). Claims of this kind open a possible new field of discussion, devoted to the novelty of what is at stake, to its epistemological implications and its exact relations with the traditional semantic and methodological forms of incommensurability. My paper and Sankey's comment on it address several aspects of these questions.

9. A NEW INTEREST IN SCIENTIFIC STABILITY, CONTINUITY AND CUMULATIVITY

The shift of focus, in the philosophy of science, from theory to practice, is not the only change that deserves mention. If it was discontinuities, ruptures, revolutions and incommensurabilities that previously struck philosophers of science, in recent times,

[3] See my paper for all the references.

an increasing number of authors have insisted on the need to attend to important continuities and stabilities that science manifests. Ian Hacking (1999) expressed the point in a striking manner, in his book, *The Social Construction of What?*: "The old idea that sciences are cumulative may reign once more. Between 1962 (when Kuhn published the Structure) and the late 1980s, the problem for philosophers of science was to understand revolution. Now, the problem is to understand stability."

This is the reason behind the third element of the subtitle of this book, "stabilities", side by side with "ruptures" and "incommensurabilities". In the present volume, this renewed interest in scientific stability is manifest in the contributions of several authors: Hartmann, who presents a repertory of different types of continuities or "correspondences" between theories; d'Espagnat, who argues that science is cumulative since predictions of observations and recipes for experimental actions are themselves cumulative; Harré, who stresses ontological continuity as a policy of research; Nola and Clarke, who argue for denotational and referential continuity.

10. THE SEARCH FOR CONTINUITY AS AN OLD PHILOSOPHICAL STRATEGY AGAINST THE INCOMMENSURABILITY THESIS

With respect to stability, continuity and cumulativity, we can note that, from the outset, it has been a common strategy to look for stable items in the history of science in order to bypass the alleged undesirable consequences of incommensurability, or more exactly, to confine the ruptures to a certain level of science, and to claim that, at another level, we find trans-paradigmatic invariants that can be identified as the carriers of a cumulative scientific progress. This has been the strategy of many realists with referential invariants (Nola's research program can be seen as a new attempt of this kind); the strategy of some rationalists with methodological invariants; as well as the strategy of some instrumentalists and pragmatists with predictive and performative invariants.

Along related lines, and in the context of the discussion of the semantic incommensurability of theories, it has been early recognized that in order for semantic incommensurability to create an epistemological problem, the semantic ruptures have to arise between *competing* theories, that is, between theories that must have something significant in common (otherwise they would not be in a position to compete). The claim that physics and psychoanalysis are two incommensurable theories, for example, does not seem to have any epistemologically damaging implications, since such theories are not supposed to be rival theories. Competing theories are theories that must have some important common points. They must notably take some common phenomena as relevant to their field, which means that they must be connectible, at a certain level, with a set of common "basic statements" (in Popper's sense) and hence with some common semantic categories (Soler, 2003). This is an in-principle minimal requirement, to which it has been added that in fact, the prototypical pairs of contemporary incommensurable theories share more than that: they always share some theoretical problems and theoretical concepts. In other words, the incommensurable areas are circumscribed, the incommensurability is only local (as Kuhn (2000a) explicitly recognized in 1983) – although it can be *more or less* local. In brief, semantic incommensurability cannot be understood in too extreme a sense. We cannot understand "incommensurable" too literally: philosophically interesting semantic incommensurability presupposes some common points between theories. Hence there must be a minimum of continuity in the history of science, even when a revolution occurs. This has sometimes appeared paradoxical, but this is one way to understand Kuhn's "essential tension" between radical novelty and the need of a certain conservation of the tradition. A new theory would not be recognized as relevant (and a fortiori as better) to an ancient one, if the new one shared nothing in common with the old and could not be connected with anything of the old tradition.

This problem arises in the context of my own paper and Howard Sankey's commentary. I extend the "quasi-analytic" association between rivalry and incommensurability verdicts to the case of experimental incommensurability. Sankey points to difficulties that have never been fully resolved in the discussion of semantic incommensurability, in particular, the problem of how assertions stated in semantically incommensurable languages can be rivals, while the absence of shared meaning entails that they are unable to enter into a relation of conflict. He argues that structurally analogous difficulties arise in relation to the experimental incommensurability proposed by Hacking and Pickering. For it remains to be explained how exactly different sets of instruments or experimental practices are able to enter into a relation of rivalry with one another.

11. PHILOSOPHICAL ISSUES ASSOCIATED WITH THE THESIS OF SCIENTIFIC CONTINUITY

With respect to scientific stability, the philosophical task is to identify what exactly has remained stable and grown cumulatively in the history of science, to evaluate the significance of the stable core that can be isolated in retrospect, and to propose an explanation of the observed stability. This task, originally focused on theories, has been directed as well toward scientific practices. With respect to this task, we can identify the following difficulties.

The first difficulty is to identify the alleged stable core. This is not so easy, especially if we take into account the partially holistic nature of science and knowledge, and consider that the answer will be heavily dependent on what we take to be the origin of genuine science. As Hartmann stresses, there are always multiple possible relations between theories and multiple correspondence strategies. Even if continuity is real, continuity appraisal is ambiguous. It depends on an evaluation of the significance of the multiples similarities and differences that always exist. Where some philosophers see continuities (as does Nola in the case of the history of the electron), others might see ruptures (see, e.g., Kuhn's (2000b) discussion of Putnam on the case of water).

The second difficulty is to evaluate the significance of the isolated stable core: it is always possible, retrospectively, to find common points between two stable stages of scientific development, otherwise we would not be ready to consider both of them as "scientific" and we would not be inclined to see them as two entities that call for a comparison. But it remains to discuss if the stable layer that is recognized to be shared by two stages of scientific development is sufficiently representative of the old science, or is nothing more than a minor, a posteriori artificially extracted part, of the beliefs and techniques that were efficiently at work in the old scientific practices.

Especially significant in this respect is the dialogue between Bernard d'Espagnat and Marcel Weber. D'Espagnat sees a significant continuity where Weber, who represents Kuhn's point of view, diagnoses a radical discontinuity. The dialogue illustrates how the description of the same historical sequence may differ, depending on what matters to the eyes of the analyst. D'Espagnat, who represents the point of view of the physicist, argues that cumulativity at the level of predictions is sufficient to conclude to the cumulativity of science. Weber, who represents the point of view of the historian and is anxious to take into account the experience of practitioners at the time, disagrees. According to him, for a scientist who experienced a revolution and came to adhere to a new paradigm, the change is radical. It is artificial and misleading to isolate a set of stable predictions from the other elements that went with them at the time and gave them sense.

The third difficulty lies in the explanation of the stability. Are we going to explain it on the basis of constraints coming from the supposedly unique invariant referent of the scientific study, in the spirit of scientific realists? Or are we going to explain it on the basis of more or less universal human categories, schema and commitments, for example commitment to conservatism, economy, convenience or the like (in the spirit of Suárez's suggestions in his comment on Harré's paper)? Or are we going to introduce into the picture elements such as the necessity for scientists to manifest their affiliation to the tradition in order to promote their careers? Or are we going to sustain any possible combinations of these options? Hartmann addresses this problem and looks for a solution that includes both "internal" and "external" factors. He attempts to explain scientific continuity in the framework of what he calls an "epistemological coherentism". According to this conception, scientists would strive for continuity, and should strive for continuity, because continuity increases coherence (leads to a better "hanging-together") and because "coherence is truth-conducive". Hartmann tries to build a coherence measure in a Bayesian framework.

The fourth difficulty lies in the legitimacy of projecting the present stability in the future and in the status ascribed to present stable results: can we admit that what we take to be as the common stable core of science now, will remain, for the most part at least, the stable core of future science, as realists and inevitabilists suggest? Can we consider that what is presently taken as a stable set of sound scientific results had inevitably to occur? Barberousse's paper and my own touch on these questions.

Acknowledgments

My gratitude goes first of all to Howard Sankey, co-editor of this volume, for his assistance in the preparation of this volume, and for interesting intellectual exchanges associated with this task. I am grateful to him as well for his support, his good humour, and the amount of time he spent improving my English expression.

I am grateful also to Paul Hoyningen-Huene, co-editor of this volume and co-organizer of the conference held in Nancy in June 2004, as well as to the other members of the scientific committee, Gerhard Heinzmann, Howard Sankey, Marcel Weber and Hervé Zwirn.

I would also like to thank all of the contributors to this volume and participants in the conference, both for the quality and openness of exchanges, as well as for the rigour and care with which they have undertaken detailed revisions of the papers they have contributed to this volume.

I wish to express my gratitude to the Henri Poincaré Archives (Laboratory in Philosophy and History of Science, UMR CNRS 7117), which provides researchers with not only an exceptional intellectual context but a convivial atmosphere, especially due to the personality of its Director, Gerhard Heinzmann, but also due to the solidarity of its members, which is even more effective because it is independent of any hierarchical structure.

Neither the conference nor this volume would have been realized without the financial support of a number of institutions. I wish in particular to thank the National Scientific Research Centre (Centre National de la Recherche Scientifique); the Urban Community of Grand Nancy; the General Council of Meurthe and Moselle; the Regional Council of Lorraine; the Scientific Council, Department of Philosophy and the "UFR Connaissance de l'homme" of the University of Nancy 2; the National Polytechnic Institute of Lorraine (INPL); the French Minister of Research; and the University of Nancy 1. I would also like to thank the two institutions responsible for the organization of the conference, the Center for the Philosophy and Ethics of Science of the University of Hanover and the LPHS–Henri Poincaré Archives of Nancy.

Many thanks to Alexander Bird, Ron Giere, Hania Grosbuch, Rom Harré and Tom Nickles for their help in the improvement of the English language of the final version of this manuscript, as well as to Charles Erkelens and Lucy Fleet, in the Editorial Office at Springer, for their advice, support, and patience.

Finally, thank you to Emiliano Trizio for his assistance on numerous occasions, especially for English corrections, and for his friendship.

BIBLIOGRAPHY

Cartwright, N. (1986) *How the Laws of Physics Lie*. Oxford: Clarendon Press/Oxford University Press.

Galison, P. (1997) *Image and Logic, A Material Culture of Microphysics.* Chicago, IL: University of Chicago Press.

Hacking, I. (1983) *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science.* Cambridge: Cambridge University Press.

Pickering, A. (1989) Living in the Material World: On Realism and Experimental Practice. In *The Uses of Experiment, Studies in the Natural Sciences.* Cambridge: Cambridge University Press, pp. 275–297.

Hacking, I. (1999) *The Social Construction of What ?* Cambridge, MA: Harvard University Press.

Hoyningen-Huene, P. and Sankey, H. (eds.) (2001) *Incommensurability and Related Matters.* Dordrecht, The Netherlands: Kluwer.

Kuhn (2000a) Commensurability, Comparability, Communicability. In J. Conant and J. Haugeland (eds.). *The Road Since Structure. Philosophical Essays, 1970–1993.* Chicago: University of Chicago Press, pp. 33–57.

Kuhn (2000b) Possible Worlds in History of Science. In J. Conant and J. Haugeland (eds.). *The Road Since Structure. Philosophical Essays, 1970–1993.* Chicago: University of Chicago Press, pp. 80–86.

Pickering, A. (1984) *Constructing Quarks, A Sociological History of Particle Physics.* Chicago, IL: University of Chicago Press.

Pickering, A. (1995) *The Mangle of Practice. Time, Agency and Science.* Chicago, IL: University of Chicago Press.

Soler, L. (2000) *Introduction à l'épistémologie.* Ellipses, 2000.

Soler, L. (2003) De la commune mesure phénoménale entre théories physiques dites incommensurables, *Philosophia Scientiae*, 7(2), 239–265.

Soler, L. (2004) The Incommensurability Problem: Evolution, Approaches and Recent Issues. In *Le problème de l'incommensurabilité, un demi siècle après/The incommensurability Problem, Half a Century Later. Philosophia Scientiae*, 8(1), 1–38.

Soler, L. and H. Sankay eds. (forthcoming) Are the results of our science contingent or inevitable?, *Studies in History and Philosophy of Science*.

PART 1

INCOMMENSURABILITY, AS DIFFERENCES IN QUASI-INTUITIVE COGNITIVE CAPACITIES: A TASK FOR PSYCHOLOGY?

INCOMMENSURABILITY NATURALIZED

ALEXANDER BIRD

Abstract In this paper I argue that we can understand incommensurability in a naturalistic, psychological manner. Cognitive habits can be acquired and so differ between individuals. Drawing on psychological work concerning analogical thinking and thinking with schemata, I argue that incommensurability arises between individuals with different cognitive habits and between groups with different shared cognitive habits.

Keywords Kuhn, incommensurability, world change, paradigm, exemplar, cognitive psychology, analogy, schema, model.

1. INTRODUCTION – INCOMMENSURABILITY AND QUASI-INTUITIVE COGNITIVE CAPACITIES

Incommensurability was one of the primary topics in the philosophy of science of the later twentieth century, having been introduced principally by Thomas Kuhn while also vigorously promoted by Paul Feyerabend. In *The Structure of Scientific Revolutions* (1962) Kuhn discussed two kinds of incommensurability – incommensurability of standards and incommensurability of meaning. He spent much of the rest of his career developing the latter using various ideas from the philosophy of language. Interest in such "classical", Kuhnian accounts of incommensurability has declined from its peak, largely because most philosophers hold that other developments in the philosophy of language have shown how incommensurability may be defeated or at least deflated (as is amply demonstrated in the work of Howard Sankey (1994, 1997)).[1]

Kuhn's later work on incommensurability contrasts with the tenor of much of *The Structure of Scientific Revolutions*. The latter I regard as having a strong naturalistic streak that was lost in Kuhn's subsequent discussion of incommensurability, which was much more standardly philosophical and aprioristic in character (Bird, 2002, 2005). While I agree with Sankey's criticisms of meaning incommensurability, I believe that a return to Kuhn's earlier work will provide us with the starting point for a rather different approach to the phenomenon of incommensurability. First, we may revisit the issue of incommensurability of standards that Kuhn did not develop in much detail. And, secondly, we may explore

[1] Declined but very far from evaporated. See Hoyningen-Huene and Sankey (2001) and this volume as examples of considerable continued interest in the general issue of incommensurability and in allied issues.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 21–39.

how Kuhn's early naturalistic account of the functioning of paradigms (as exemplars) may furnish a naturalistic, primarily psychological conception of incommensurability.

The key idea in what follows is that we all use in thinking various cognitive capacities and structures that have the following features: (i) they cannot be reduced to general, formal rules of reasoning (e.g., formal logic); (ii) their existence and the mechanism of their employment are typically unconscious, so that they are deployed in a manner that is akin to intuition – what I call a *semi-intuitive* manner; (iii) they are often acquired as a matter of practice and repeated exposure, so that they have the character of skills. The sorts of skill or capacity I am referring to here include: mental schemata, analogical thinking, pattern recognition, quasi-intuitive inference. As I shall describe below, these are related, and together I shall refer to them as an individual's quasi-intuitive cognitive capacities (QICCs).

The proposal of this paper is that as a result of social induction with a set of paradigms a scientist in a given field acquires a set of QICCs specific to that field. Incommensurability arises when scientists operate with different QICCs; this incommensurability is an incommensurability of standards and an incommensurability of understanding (which is not quite the same as an incommensurability of meaning, unless meaning is understood in the relatively loose sense corresponding to the intended message of a communication). Such incommensurability may arise in communication between radical and conservative scientists during a scientific revolution, in the understanding of the work of scientists from another era, and also between scientists working on similar problems but from differing background fields.

2. EXEMPLARS IN SCIENCE – AN EXAMPLE

Much cognition is habitual. These habits are acquired. This occurs in scientific thinking no less than in informal, everyday cognition. Consider the following example of a test question that might be set for physics students:

> A thin rectangular plate, *O*, of mass *m* is free to rotate about a hinge at one edge. The length of the plate from the hinge to the other edge is *l*. It is placed within a stream of fluid which exerts a pressure P in the direction of flow. Let θ be the angle between the plate and the direction of fluid flow. Write down an equation describing the motion of *O* in terms of the change of θ over time, assuming θ is small. (See *Fig. 1.*)

Many will be able to see immediately that the motion will be simple harmonic and will be able to write down the equation of motion straight away as having the form: $\theta = \theta_{max} sin(kt)$. In some cases this will be assisted by seeing an analogy between the oscillating plate and a rigid pendulum (see Fig. 2). The motion of the

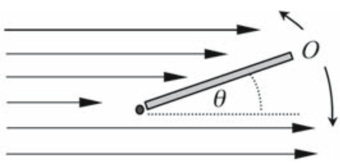


Fig. 1.

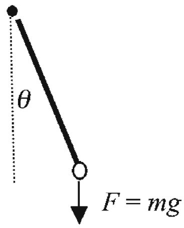


Fig. 2.

pendulum is the first and best known instance of simple harmonic motion that any student of physics gets to meet. Seeing the analogy will thus enable the student to know the correct equation of motion, or at least its form, without further ado. The analogy in this instance will be easy to spot, since the standard diagram for a pendulum is very similar to that employed in the question given, rotated through a right angle.

This example illustrates important aspects of cognition in science:

(i) The employment of paradigms. Kuhn (1962) defines paradigms as standard exemplars of problems and their solutions. In this example, the motion of the pendulum and its standard solution are a paradigm.

(ii) Analogical thinking. The use of a paradigm is essentially analogical. The analogy between the new problem of the oscillating plate and the well-known case of the pendulum shows or suggests to the student that the equation for the former will have the same form as that for the latter.

(iii) Scientific cognition is not a matter of following very general rules of rationality (where "rule" is understood in a strong way, implying that it can be followed mechanically). (This was a common background assumption among many logical empiricists, and the significance of paradigms in Kuhn's work is that they provide a refutation of the logical empiricist and positivist view of scientific cognition.)

(iv) Learned, quasi-intuitive inference patterns and associations: the ability to answer immediately or via the pendulum analogy is quasi-intuitive. It is like intuition in that it is non-reflective or only partially reflective, being a more-or-less direct inference (rather than one mediated by the sort of rules mentioned in (iii)). But it is unlike intuition in that it is learned, acquired by training with paradigms/exemplars (e.g., many student exercises are of just this sort).

(v) Cognitive processes can be acquired habits. The quasi-intuitive inferences mentioned in (iv) are made as a matter of habit. The habits are acquired as a result of repetitive exposure and practice. For example, what may start out as a conscious, sequential activity of reasoning, eventually becomes a one-step quasi-intuitive inference. This does *not* mean that the same process of sequential reasoning takes place unconsciously. Rather the habit replaces the conscious reasoning process.

3. SCHEMATA AND ANALOGIES IN SCIENTIFIC THINKING

I take the above to be ubiquitous features of scientific cognition (and indeed of cognition in general). But the illustration given by one rather low-level example is not sufficient to show this. Nor does it show how quasi-intuitive inferences dovetail with the undeniable existence of sequential, conscious ratiocination in science. Even so, other work confirms the claim that these features are indeed widespread in science.

First of all, it is clear that cognitive and inferential habits are a general feature of human psychology. This has been recognized since Hume and before, and has recently been the subject of more systematic study. For example, educational psychologists, employing an idea originating with Piaget have developed schema theory, according to which thinkers employ *schemata* which link certain patterns of stimuli with responses in such a way that does not require the conscious attention of the subject (Schaller, 1982; Anderson, 1984; Stein and Trabasso, 1982; D'Andrade, 1995). (*Scripts* are also often discussed, which are schemata that apply to behaviour.) A schema encapsulates an individual's knowledge or experience, and is typically acquired through repeated use of a certain pattern of inference. Although schemata have not been discussed especially in relation to science, if schemata, or something like them (e.g., mental models, which can be regarded as a more general kind than schemata) are ubiquitous in thinking outside science, it is reasonable to suppose that they play a role in scientific thinking also. Schemata would correspond the to the quasi-intuitive inferences I have mentioned above. One may then hypothesize that the function of Kuhnian exemplars is the inculcation of scientific schemata. It is notable that while many schemata are individual (and some may be innate), many are also shared and culturally acquired, as are Kuhnian exemplars.

Exemplars also function analogically. Historical work, most notably that of Mary Hesse (1966), has shown how prevalent analogy is in the history of science, while contemporary psychological work by Kevin Dunbar (1996, 1999) confirms that analogical thinking is an important feature of modern scientific thinking.[2] Rutherford's analogy between the solar system and the structure of the atom is a well-known example. Dunbar studied scientists at work in four microbiological laboratories. He found that use of analogical reasoning was very common. Distant analogies, such as Rutherford's were unusual, but close analogies were frequently used in the design of experiments (very close – e.g., one gene on the HIV virus to another on that virus) and the formulation of hypotheses (moderately close – e.g., the HIV virus and the Ebola virus). Analogy is particularly helpful in explaining unexpected results: again, local analogies are employed first in coming to a methodological explanation; if a series of unexpected results survives methodological changes, then scientists will use more distant analogies in the construction of models to explain their findings. Analogical reasoning has been modelled successfully with computer simulations; of particular interest is the modelling of the analogical use of old problem solutions in solving new problems, known as *case-based reasoning* (see Leake, 1998), which neatly describes the particular use

[2] See also Holyoak and Thagard (1995) on analogical reasoning, particularly in science. But note that Gentner and Jeziorski (1993) also argue that the degree to which analogy is used differs through history, and that medieval scientists used metaphor to a greater extent than true analogy, relative to more recent scientists.

of analogy described in the example of the hinged plate and the pendulum. While case-based reasoning is most conspicuous in medicine and most especially in psychoanalysis,[3] Dunbar's work as well as that of Hesse and of the historians of science shows how it is a ubiquitous feature of scientific cognition.[4]

Kuhn stresses the role of exemplars in forging relations of similarity among problems. Research (Chi et al., 1981) confirms that training with scientific problems induces new similarity spaces on problems and their solutions. Thus novices (undergraduates) categorize problems differently from experts (graduate students). The former employ superficial criteria (e.g., similarity of diagrams and keywords – i.e., categorizing problems as rotating disk versus block and pulley problems) whereas the latter categorize the problems according to the underlying physical principles involved (conservation of energy versus force problems). What role does the acquired similarity space play in scientific problem solving? Clearly, if one is able to categorize problems according to their underlying physical principles then one is in a good position to start focusing on the equations and laws that may be used in solving the problem. In the light of the forgoing discussion of analogies and schemata, one key function for a similarity space is the selection of an appropriate schema or the bringing to mind of an appropriate analogy. More generally we can see the possession of a similarity space of this kind as an instance of a pattern-recognitional capacity. What are seen as similar are certain structures or patterns. When we use a schema or analogy we see that the target has the structure appropriate for the application of the schema or one that it shares with the source analogue. It is possible to see most or even all of scientific thinking as involving *au fond* pattern-recognitional capacities – Howard Margolis's (1987) account of scientific cognition is just that.[5]

Analogical thinking and thinking in schemata are not *prima facie* the same. Schemata are abstract structures with place holders for various possible inputs and outputs. Analogies are more concrete. The source analogue (e.g., the solar system in the Rutherford example) is complete rather than possessing place-holders. Nonetheless there are clear similarities and relations between the analogues and schemata. Any interesting use of analogy involves abstracting salient features of the source analogy. Thus the structure abstracted from an analogue will have the character of a schema. The nodes in such an abstracted structure are now place-holders. If we take Napoleon's invasion of Russia in 1812 as an analogy for Hitler's invasion of the Soviet Union in 1941, then we are abstracting "*x* invaded Russia/Soviet Union" with *x* as a place-holder that may be taken by either Napoleon or Hitler. Thus we can see that use of an analogy may lead to the creation of a schema.[6] We start by thinking analogically, but

[3] See Forrester (1996) on thinking in cases in and beyond psychoanalysis, linking it to Kuhnian themes.

[4] Nickles (2003a) also relates Leake's case-based reasoning to science and to Kuhnian exemplars. Nersessian (2001, 2003) discusses (in)commensurability in relation to model-based reasoning. My approach is sympathetic to Nersessian's "cognitive-historical" approach. Unlike Nersessian, I am keen to explore, as far as possible, the contribution of cognitive science in a manner that is independent of issues of concept formation and conceptual change (cf. footnote 12). That said, Nersessian (2001, p. 275) does distance herself from the view that these matters are to be understood in purely linguistic terms.

[5] This we may in turn think of as being realized by connectionist systems in the brain. For how this might work in the scientific case, see Bird (2005); cf. Nersessian (2003, p. 185).

[6] Cf. Gick and Holyoak (1983) and Holyoak and Thagard (1997, p. 35) on analogy inducing a schema. Gick and Holyoak note that problem solving with analogies is more effective when achieved by reasoning via an induced schema as against reasoning directly from an analogue.

repeated use of a certain analogue will in due course allow us to think directly with the schema it gives rise to. Alternatively, we may regard the source analogue as a concrete vehicle for the schema. The analogue is concrete, but its use is by reference only to the abstract, schematic structure it possesses.

A further difference between schema and analogy is that the possession and operation of the former is typically unconscious while the use of analogy is conscious. Nonetheless, the similarities are more significant. First, there are unconscious elements in the operation of an analogy: the selection of an appropriate analogy may come immediately as may the forging of the structural links and the drawing of conclusions from them. In the example given above, the problem solver may immediately think of the pendulum as a good analogy and having done so immediately see what implications that has for the equation of motion of the hinged plate. Secondly, in both cases the reasoning and inferential processes involved are not algorithmic – for someone with an appropriate schema or who is adept at using analogies in the way described, the deployment of the schema/analogy is intuitive. Of course, possession of a schema or facility with an analogy, are acquired capacities, a matter of learning or training, and will often involve repetition and practice. So the "intuition" is learned. That may sound like an oxymoron (so prompting my use of "*quasi*-intuitive"), but nevertheless captures the entirely coherent idea that cognitive habits can be acquired. It is the same phenomenon for which we use the term "second nature". First nature we are born with, second nature is acquired. But both feel natural and intuitive, rather than forced or the product of consciously following a rule or explicit ratiocination. Second nature covers a wide range of acquired habits and skills such as the fingering technique of an accomplished pianist; for our purposes the habits of relevance are cognitive ones, and in particular the quasi-intuitive inferences encapsulated in schemata and analogies.

Altogether, I shall refer to schemata, analogical thinking, quasi-intuitive inferences, similarity spaces, and pattern recognition as *quasi-intuitive cognitive capacities* (QICCs). In what follows I shall argue that if as claimed QICCs, acquired as a result of training with exemplars, are a common feature of scientific thinking, then we should expect science to be subject to a certain kind of incommensurability. I shall then suggest that it is this phenomenon to which Kuhn was referring in *The Structure of Scientific Revolutions*. This would permit a naturalistic understanding of incommensurability. It is naturalistic because the features described above are not *a priori* characteristics of scientific cognition; to appreciate their existence and to understand their nature requires empirical investigation of science. To that extent Kuhn's earlier historically founded account of scientific change also counts as naturalistic. And it is noteworthy that Kuhn remarked in 1991 "many of the most central conclusions we drew from the historical record can be derived instead from first principles," (1992, p. 10) thereby insisting that even his own earlier historical naturalism was unnecessary. However, the naturalism of Kuhn's interest in psychology and of the approach taken here goes further, in that it draws upon the *scientific* investigation of science and scientific cognition. It is true that when it comes to psychology, the boundaries between the *a priori* and the empirical and between the empirical-but-not-scientific and the scientific may be quite unclear.

For it may be that some proportion of folk psychological knowledge is innate, and it is certainly true that much psychological insight may be acquired without science simply through reflective acquaintance with the variety of persons about us and through reflection on our own cases. While Kuhn's basic insight may not itself be a product of scientific research, he was certainly willing in *The Structure of Scientific Revolutions* to draw upon the results of scientific experimental psychology, be it Gestalt psychology or the playing card experiments of Bruner and Postman. Furthermore, there are few insights of subjective or folk psychology that cannot be improved and put on a firmer epistemic footing by scientific psychology. The proposals as regards incommensurability I put forward here do not presume to constitute a scientific investigation of that phenomenon. I do believe that they have some plausibility by appeal to our low level psychological knowledge. But they do also draw upon the scientific investigations referred to above and, more importantly, they are programmatic in that they suggest an approach that stands open to scientific development and confirmation or refutation.

According to my proposal the features of science that permit incommensurability exist in virtue of the nature of human psychology (principally the QICCs). The nature of incommensurability must be regarded as fundamentally psychological. This contrasts with more explicitly *a priori* (principally semantic) accounts of incommensurability given by many philosophers, including Kuhn himself in later life, which draw upon the philosophy of language broadly construed. Insofar as incommensurability is to be found between groups of scientists, there is also a sociological dimension. But the basic component is still psychological. It is the fact that groups of scientists are all trained with the same or similar examples and the fact that familiarity with those examples and a grasp of the quasi-intuitive inference patterns they engender are pre-conditions of membership of the group that explains why incommensurability emerges as a social phenomenon. (In a similar way language is a social phenomenon. But what makes it possible are features of individual psychology.)

4. INCOMMENSURABILITY OF STANDARDS

Incommensurability, Kuhn tells us, is a matter of a lack of common standards. It is easy to see how incommensurability may arise from the psychological features of scientific cognition outlined above. Let us assume for a moment that the QICCs discussed also encompass the capacity for making judgments of correctness. For example, a judgment of correctness may be made on the basis of the use of some schema induced by experience with a particular set of exemplars, or on the basis of perceived similarity to such exemplars. If that is so, then scientist A and scientist B who have been trained with different sets of exemplars, and so have different schemata or similarity spaces may not be able to come to the same judgment over the correctness of some putative scientific problem-solution. What goes for individual scientists goes also for groups of scientists, where the two groups differ in their exposure to exemplars.

In this section I shall consider the question, can the account of QICCs be extended to include judgments of correctness? In the next section I shall extend the account to encompass incommensurability of *understanding*, by which I mean the phenomenon of incomprehension that Kuhn identifies between adherents of different scientific schools or sides in a revolutionary dispute.

Even if cognition often involves QICCs, it does not immediately follow that judgments of correctness involve a similar mechanism. One might invoke the familiar context of discovery versus context of justification distinction. Employing this one could claim that QICCs may lead one to an appropriate answer to a problem, but that one need not use that same mechanism in showing that the answer is a good one. For example, in formal logical systems, such as first-order Peano arithmetic, there is no algorithm for generating a proof of any given proposition. But given any putative proof there is an algorithm for checking whether or not it is a proof. Mathematician A and mathematician B may come at their proofs in different ways, but they will be able to agree on the correctness of a proposed proof. Similarly two scientists may produce their theories in different ways as a result of differences in training with paradigms, but that does not preclude their being able to make similar judgments of the correctness of some given theory.

Such a response underestimates the role of exemplars in cognition, of the QICCs they induce. Let us consider the mathematical case. It is in fact both rare and very difficult in mathematics to reduce the proofs that mathematicians do offer to a logical form that may be checked algorithmically. For a start, the possibility exists, even in principle, only when the mathematical field in question has been formally axiomatized, and that is a relatively modern development and is still not ubiquitous. It is worth noting that Euclid's axiomatization of geometry fails in this respect, not only because it is not formalized within a logical system, but also because it too depends for its operation on quasi-intuitive inferences. Indeed, much of the impetus behind the programme of rigor, formalization, and the development of logic in the nineteenth century arose from the realization that even Euclid's proofs rested upon intuitive assumptions, in many cases generated by a visualization of the geometrical problem in question. Thus what standardly passes as proof, even today, is not something that can be algorithmically checked. If it were, then the checking of proofs would not be the difficult and laborious process that it is. Moreover, mathematicians are not interested in proof merely as a means of validating belief in theorems. Rather the proofs themselves furnish understanding and insight – at least a good proof does – which is why mathematicians may seek new proofs of old theorems. But that understanding cannot be gained by looking at the proof at the level of formalized steps, but only by having an overview of the proofs, and that will involve seeing how the proof works at a level such that when "A follows from B" the relationship between A and B is of the quasi-intuitive inference kind rather than the formal logical step kind.

Thus even in the paradigm case of a discipline where contexts of discovery and justification can come apart, it seems that QICCs play a part in the context of justification as well as that of discovery. If that is true of mathematics, how much more true is it of empirical science? What has been said about mathematics is true also of the mathematical components of a scientific paper. Furthermore, the relationship of confirmation is one that cannot be formalized. More generally, and of particular importance, is the fact that any scientific paper rests on a mountain of unstated assumptions and background knowledge. These are assumptions concerning the mathematics employed, theoretical background knowledge common to all working in the field, and practical background knowledge concerning the experimental apparatus and other aspects of the experiment. These are all encoded in the schemata that a trained scientist (quasi-)intuitively employs.

Thus a difference in schemata, standard analogies, similarity spaces, and hence the sets of quasi-intuitive inferences scientists will possess and employ, may be expected to lead scientists from different traditions to make different judgments of correctness.

5. INCOMMENSURABILITY OF UNDERSTANDING

Although Kuhn makes clear that incommensurability means the lack of common standards of evaluation, he nonetheless emphasizes the connection between incommensurability and matters of meaning, and indeed the search for an account of incommensurability in terms of meaning was the focus of much of his philosophical writing in the later part of his career.

In the last section I mentioned the importance of background knowledge and tacit assumptions. These are encoded in schemata, and since background beliefs influence our judgments, we may expect that those with different sets of background beliefs may make different judgments. However the influence of tacit assumptions is not limited to the making of judgments but also extends to communication. For example, if subject A asserts "P therefore Q" where the inference from P to Q is a quasi-intuitive one, mediated by a schema that subject B does not possess, then B may not be in a position to understand A's reasoning. In this case the literal meaning of "P therefore Q" is comprehensible. But communication goes beyond mere literal meaning.

Almost all communication depends on shared tacit assumptions. To employ a homely example, a cook's recipe may state "break two eggs into a bowl; stir in 200 g sugar and 100 g butter". The recipe does not tell one to exclude the egg shells from the bowl; that's just obvious, at least to a cook, as is the fact that the sugar should not be in cubes or the butter frozen. We all make tacit assumptions of these kinds all the time in our conversations with one another. And this is no less true in science. Indeed it is more true in science because there is so much background knowledge that it simply would not be possible to refer to it all explicitly. The function of schemata is to encode such information. Thus a cook might have an EGG-IN-COOKING schema that tacitly encodes the instruction to keep the yolk and white but discard the shell (unless explicitly instructed otherwise). Similarly scientists will have schemata that encode information about the use and purpose of standard pieces of equipment, schemata for various common kinds of inference patterns, and likewise schemata for encoding all sorts of other knowledge. The function of a scientific education is, in good part, to furnish young scientists with the schemata required to participate in their speciality. Since all scientists in the given field share the same schemata, communication between them is schemata-dependent. Even a schoolboy can say and understand, "the litmus paper turned red, therefore the solution is acidic" without having to enunciate or hear the major premise, "acids turn litmus paper red", since the latter is part of the LITMUS schema, one of the earliest acquired in a science education. That inference exemplifies a simple quasi-intuitive inference. A slightly more sophisticated instance would be the familiar pattern found in reasoning concerning simple harmonic motion, as may be applied to the problem with which this paper starts: (i) $d^2\theta/dt^2 = -k^2 \sin \theta$, therefore (ii) $d^2\theta/dt^2 \approx -k^2\theta$ for small θ, therefore (iii) $\theta \approx a \sin kt + c$ for small θ. Those inferences will not be comprehensible as rational inferences by someone who lacks the training and hence schemata common to mathematicians, physicists and engineers.

The role of tacit assumptions is not limited to underpinning (enthymatically licencing) the rationality of inferences that would otherwise look incomplete or irrational. For those assumptions also play a role in what a speaker intends to communicate. In the cooking example, the writer's intention is that the whiteshell is excluded. Linguists have examined the role of tacit assumptions in intended meaning, for example, in rendering unequivocal a case of potential ambiguity (cf. Carston, 2002; Sperber and Wilson, 1995). But as they also note, such assumptions play a role not only in mediating between the words uttered and their literal meaning but also between literal meaning and the speaker's intended message, what the speaker means to say in a broader sense.

In many cases we can fairly easily retrieve the tacit assumptions and make them explicit, but in other cases it is far from easy.[7] A tacit assumption – a proposition included in a schema – will not always be one that was once explicit to the subject. There are many tacit assumptions with which we are born and which constitute the basis of folk physics and folk psychology as well as what one might call folk metaphysics (e.g., assumptions about causation – experiments suggest, for example, that small children expect causation to involve contiguity in time and in space). Such assumptions are difficult to retrieve because they have never been explicit. They have been hard-wired from the beginning. They ground fully intuitive inferences. Quasi-intuitive inferences occur as a result of a habit of using an explicit inference pattern repeatedly, such that, one can reasonably assume, the relevant neural pathways are developed in a way that is analogous to those that were hard-wired from the start. In such cases the need for the explicit assumption falls away. But note that the quasi-intuitive inference form may be passed on to other subjects without the tacit assumption ever having been explicit for *them*. Hence they will not be able to use memory as a tool for retrieving those assumptions. We often accept the truth of propositions on the basis of testimony; similarly we accept the reliability of inferences on the same basis. In the former case we lack the evidence for the proposition in question; in the latter case we may be ignorant of the intermediate propositions that mediate the inference pattern. In such a case "retrieval" will not be so much a matter of finding a tacit assumption that the subject once possessed explicitly (for she never did) but will be rather more like a rational reconstruction of the inference pattern. In other cases the inference, although not innate, is one that *feels* natural.[8] The fact that tacit assumptions may be difficult to retrieve or reveal is

[7] By "retrieving" I mean coming to know that such-and-such is an assumption that one makes. One common case of the difficulty of retrieval is found in communicating with small children. Parents have to retrieve many common tacit assumption that adults employ but children do not.

[8] For example many of us find the inferences made in simple calculus compelling, such as: let AB be a chord on the curve C; then $(y_B - y_A)/(x_B - x_A)$ is the slope of AB. As B approaches A, $(y_B - y_A)/(x_B - x_A)$ approaches m. Therefore the slope of the tangent to C at A is m. The justification for this inference provided by Newton implied that when A and B coincide $(y_B - y_A)/(x_B - x_A) = m$. But as Berkeley pointed out $(y_B - y_A)/(x_B - x_A)$ is just 0/0 when A = B, which is undefined. Mathematics had to wait until Cauchy's rigorous definition of a limit before a satisfactory arithmetical justification was given of an inference form that everybody held to be valid. What exactly makes this inference feel acceptable is

significant for an account of incommensurability, since it explains the reticence and incomprehension many feel when faced with a revolutionary proposal. Understanding and accepting the proposal requires the jettisoning of tacit assumptions, which will not be easy if they are difficult to retrieve or reconstruct.

6. EVIDENCE FOR PSYCHOLOGICAL INCOMMENSURABILITY

What ought to be the best test of this explanation of incommensurability would be to use it to account for paradigmatic cases of incommensurability, such as incommensurability in disputes between adherents of old and new "disciplinary matrices" (to use the term Kuhn used to capture a broader notion of paradigm, encompassing an array of commitments of a scientist or groups of scientists). This is, however, not as straightforward as it might seem it is not obvious that there is some clear phenomenon of incommensurability with paradigmatic cases we can point to. The very phenomenon is disputed. Kuhn did point to a number of alleged instances in *The Structure of Scientific Revolutions*, such as the Proust–Berthollet and the Priestley–Lavoisier disputes. But their treatment was not sufficiently detailed that we can unequivocally identify incommensurability as opposed simply to a scientific dispute.

For we should not immediately infer incommensurability from disagreement, since disagreement, even among rational individuals, does not show difference in standards or in understanding. To regard all rational disagreement as betokening incommensurability would be to trivialize the latter notion. The pages of *New Scientist* and *Scientific American* are full of examples of scientific disagreements. But to classify all of these as exemplifying incommensurability would be to render the latter concept of little interest. How should we distinguish disputes that involve incommensurability from "mere" scientific disagreements? In most ordinary scientific disagreements the disputants will genuinely understand each other and will typically acknowledge the relevance of one another's arguments. What will differ between them will be the weight they attach to the various considerations. In more serious disputes, those displaying incommensurability, we should expect to find the parties differing over even what counts as a relevant consideration, and in some cases to find incomprehension between the parties. If so, then incommensurability can be explained by the current account, as we have seen above. For the relevance of potential evidence depends on background knowledge, encoded in schemata, and so a disagreement over the very relevance of evidence may be readily explained by a difference in schemata. Likewise schemata are required for comprehension, and so a lack of understanding will also be explained by a lack of shared relevant schemata.

One problem is that it is natural to think of incommensurability as a symmetric concept: if A is unable to judge the science of B, thanks to incommensurability, then B is unable to judge the science of A for the same reason. While that may be the case in

not entirely clear. It may be some general mathematical sense that we acquire in learning mathematics; it may be partly the rhetoric of the way we are taught; it may be that we just learn the technique and familiarity makes it feel natural. In any case, the justification for the inference is not something we can retrieve but is learned later as a rational reconstruction of what we already accept.

certain instances, we should not expect to find it in revolutionary disputes between the adherents of the old disciplinary matrix and supporters of a revolutionary replacement. For the latter will typically have been educated within the old paradigm and will be well placed to understand and assess what it is they are trying to replace. Nonetheless, we might expect the conservatives not to regard the radicals' evidence as relevant or to fail to understand the radicals' proposals. The radicals will understand the conservatives and while they may also reject their evidence, they will at least know where it originates and why the conservatives regard it as relevant.

The classic case that does seem to exemplify incommensurability of this sort is the dispute between Galileo and the philosophers in the early seventeenth century. Here it seems that the relevance of experimental evidence versus the authority of Aristotle was in dispute, and that Galileo's mechanical theories were not understood by many of his critics. It is implausible to suppose that incommensurability arose because Galileo's assumptions were tacit, hidden in his schemata. As a radical his premises would be newly developed and would be explicit.[9] Nor, for the reasons mentioned above, would it be the case that Galileo would not be aware of the tacit assumptions being made by his opponents. Rather, in this case, it must be that the schemata employed by the conservatives *prevent* them from accepting the new proposal, even when the proposal and the evidence for it are presented in as transparent and assumption-free a manner as possible. It may be that the conservative schemata make it impossible to see the relevance of the radicals' evidence. But more often those schemata will make acknowledged facts into evidence *against* the new proposal, in some cases even to the degree of making the new proposal seem senseless. Thus the conservatives see the relative order of things on Earth as a sign that it does not move or rotate (the Earth in motion would cause widespread chaos and destruction). In particular they argued that an object in free fall should not appear to fall directly downwards, but should seem to follow an arc as it is left behind by the rotating Earth. In another case they possessed a schema for falling bodies that includes the proposition that heavier bodies fall with a greater acceleration than lighter ones, which would be a schema that we all acquire naturally but also is reinforced by Aristotelian doctrine. Thus a major obstacle to Galileo's dialectical success was the need to dismantle his opponents' schemata that were in many cases deeply entrenched. But because schemata are entrenched, their existence and function is unknown to their possessors, and their contents thus difficult for their possessors to retrieve, they cannot be removed simply by the amassing of evidence against the propositions they encode. Thus the need for more "psychological" methods of engagement, for example the use of rhetorical techniques (much commented upon by historians) and thought experiments. The purpose of a thought experiment, I surmise, is to assist in dislodging a schema by making explicit its internal contradictions or its inconsistency with some undisputed fact.

In the above I emphasized the role of the conservatives' schemata in preventing acceptance of a radical view, rather than the existence of a radical schema they do not possess. However the latter can play a significant role in other kinds of case of

[9] This is not to deny that he would have developed schemata in his own thinking. Rather he would be able to retrieve the tacit premises easily and would have a reason to make them explicit.

incommensurability, and was not absent from Galileo's case. One of the factors in Galileo's dispute was that he was a mathematician whereas his opponents were not. Consequently Galileo was inclined to use forms of argument that the latter thought inappropriate and which they probably did not fully understand. Incommensurability may arise, and may be symmetrical, when a given problem is tackled by scientists coming from different disciplinary backgrounds. The backgrounds may be perfectly consistent with one another but the nature of their differing schemata is such that participants in a debate find it difficult to comprehend one another's approaches. This may partly explain the phenomenon of revolutions brought about by outsiders. The outsider's new schemata and background knowledge, as well as lack of the constraints provided by the disciplines extant schemata may allow her to see potential solutions to a problem that the discipline has hitherto been unable to see. At the same time, it may be difficult for the newcomer to persuade the current practitioners of the discipline of the value of the new solution (quite apart from the problem of disciplinary jealousies).[10] More generally incommensurability may arise when different disciplines converge on a similar set of problems. For example, an emerging area of research exists at the physics-biology interface. But part of the problem in making progress is the lack of a common approach between biologists and physicists. Thus whereas biologists are likely to take a detailed biochemical approach to understanding, say, blood or the behaviour of the components of a call, physicists will apply quite general mathematical models from soft matter theory or complex fluid theory. At a grosser level of description, physicists are theory-friendly, whereas biologists tend not to be.

As it stands evidence concerning the nature of incommensurability in actual scientific practice is largely anecdotal. Even so, what there is tends to support the view that insofar as it does exist, incommensurability may be understood using the psychological tools described above.

7. WORLD-CHANGES AND THE PHENOMENOLOGY OF INCOMMENSURABILITY

Kuhn's incommensurability thesis is closely related to his claim that as a result of a revolution a scientist's world changes. The latter has attracted attention because of its constructivist sounding content. But it is clear that Kuhn has no simplistic constructivist intent. However, what exactly Kuhn did have in mind is less clear. I propose that Kuhn's usage is only a minor extension, if at all, of a perfectly common English metaphor of world and world-change. We often talk about someone's world changing as a consequence of a major event, as we also talk about someone's world, or the world of the poet, professional footballer, etc. to refer to their milieu and the sorts of activities they engage

[10] A recent example may be Luis and Walter Alvarez' meteor impact theory of the K–T extinction. Neither was a palaeontologist – Luis was a physicist and Walter is a geologist. Until their proposal most palaeontologists were gradualists who tended to favour evolutionary or ecological explanations. Nonetheless, the Alvarez team were not complete outsiders – most palaeontologists have a background in geology, and the palaeontologist Dale Russell had already suggested in the 1960s that the dinosaurs might have been killed off by radiation from an exploding supernova.

in. For this to be extended satisfactorily to science as an account of Kuhn's notion of world, we need to show that (i) incommensurability and difference of world relate to one another, so that, roughly, scientists experience incommensurability when they are operating in different worlds; (ii) worlds have a high degree of stability and are resistant to change, but may undergo the dramatic changes associated with revolutions; (iii) that worlds have a characteristic phenomenology, in that world-change may be likened to a Gestalt switch and encountering a different world leads to a particular sense of unfamiliarity and bafflement.

The proposal is then that a scientist's world is his or her scientist's disciplinary matrix – the constellation of professional relations and commitments, instrumental and theoretical techniques, disciplinary language, and, above all, the discipline's key exemplars. And in the light of the discussion hitherto, I shall highlight an especially important and psychologically salient component of a scientist's world. This is the set of schemata, analogies, similarity spaces, pattern-recognitional capacities, and quasi-intuitive inferences that govern a scientist's response to scientific problems. Nonetheless, I do not think that a world is necessarily limited to the latter (as I suggested in Bird, 2005). I think that the key to a world is principally *entrenchment*, the difficulty with which a belief, commitment, or practice can be changed. And perfectly conscious, theoretical commitments can be among these as well as tacit assumptions. Something may be entrenched because it is unconscious and difficult to retrieve or it may be entrenched because of its central role in informing a wide range of activities. That said, the distinction is not a sharp one, not least because a belief may be, for the same scientist, a conscious theoretical assertion in one context and a tacit assumption or element of a schema in another. Thus a scientist may have an explicit commitment to, say Newton's laws of motion, or the theory of evolution through natural selection and will often mention and explain those commitments in their work. At the same time, those commitments will also be reflected in the models and analogies the scientists use and in the schemata they possess.

The proposal clearly meets the desiderata for a conception of world. First, world-change will tend to bring with it the possibility of incommensurability. The account of incommensurability I am promoting explains it by recourse to non-formal, implicit mechanisms of thought (analogical thinking, schemata, quasi-intuitive inference) that form a central part of a world, and so changes in the latter aspects of a world will thus give rise to possible incommensurability. As just mentioned a world-change may and typically will involve explicit commitments and that will require a reordering of many overt activities of the scientist. Many aspects of a scientist's professional life may change (the professional relationships engaged in, the equipment and techniques used, as well as the theories defended and employed). On their own these changes need not imply the possibility of incommensurability. But almost inevitably they will, for the reason given above, that central explicit commitments inform our tacit assumptions, schemata, and so forth.

Secondly, world-change correlates with revolutionary change. Indeed this is a trivial consequence of the definition of world in terms of the disciplinary matrix. More importantly, the entrenchment of central features of the disciplinary matrix explains the conservativeness of normal science and the resistance of science to change. Our unconscious commitments may be difficult to retrieve and alter and our conscious commitments may be so central to our activities that it may be difficult to make all the corollary changes required by changes to those commitments. It is also true that the professional interests of well-established scientists may give them an interest in resisting change. That point is certainly true, but is far from the only source of conservativeness. Another, in particular, is the role that our entrenched tacit commitments, as built into schemata and the like, play in making it difficult to engage in revolutionary change. Since these also give rise to incommensurability, those whose commitments are especially deeply entrenched may simply not be able to see the sense of the proposed change. Older scientists may well have a greater interest in the status quo. But they will also have their world more deeply entrenched and will be more likely to fall victim to the barrier to understanding erected by incommensurability, whereas the younger scientists may have their worlds less deeply entrenched.

Thirdly, we can attribute a particular phenomenology to world-change. Kuhn likened world-change to Gestalt shifts. As he acknowledged this analogy is misleading. But it is not altogether inappropriate. To begin with, Gestalt shifts are a matter of seeing one pattern and then another, and to the extent that scientific thinking involves pattern recognition, it is reasonable to consider changes in the patterns that one's exemplars, analogies, schemata and so forth permit, as either analogous to Gestalt shifts or perhaps even parallel instances of some general pattern recognitional capacity. Secondly, Gestalt shifts have a characteristic phenomenology. The picture now looks different. One of the problems for Kuhn in *The Structure of Scientific Revolutions* is that he took his psychological examples from perceptual psychology, as did his predecessors, such as Hanson. Thus the emphasis in the discussion of incommensurability gave the impression that a major component is perceptual – perception and thus observation are different as a result of a revolution. But this has significant consequences for science only if one shares the logical empiricist emphasis on observation-as-perception as an epistemological foundation. As a critique of logical empiricism that may arguably be effective. But as an understanding of an objective phenomenon of incommensurability it is very limiting, since many revolutions do not involve any change in perceptual experience. Nonetheless, I think that there is a more general analogue to the Gestalt shift, in which one's "perception" of the world changes. In this sense "perception" refers to one's (quasi-)intuitive responses to the world. This is an appropriate extension of the strict notion of "perception" and one which is standard in everyday English, since the information one gleans from the world is not limited simply to the sensory experience one has but includes, at least, also what one automatically and naturally infers. Indeed to allow a fairly extended process of inference fits with Kuhn's reminder that observation is quite different from perception.

Nor is it inappropriate to attach the tag "phenomenological" to such responses and inferences because even if the mechanics of the inference are tacit, its products – a belief, an action, an emotion, etc. – are parts of one's conscious experience and understanding of the world. As the phenomenological tradition since Husserl has emphasized, how we experience the world is much richer, much more imbued with meaning, than a focus on sensory experience would suggest.[11] The proposal here is that this is explained by the fact that our quasi-intuitive responses to the world add content and are, in a sense, interpretative (we see things in their theoretical and causal relations with other things).

Corresponding to the fact that a world brings with it a characteristic phenomenology is the fact that a world-change itself will have a phenomenological character. The same features of the world will now elicit different responses, a new phenomenology. What was unfamiliar, may now be familiar; what was previously significant may now be insignificant, or vice versa. Familiar things may be seen as standing in different relations with one another, and crucially for science problems may be seen as having solutions that could not have occurred to the scientist beforehand. Kuhn mentioned religious conversion in this connection, again attracting controversy. But the analogy has two worthwhile features. First is the fact that a change in disciplinary matrix cannot always be brought about by straightforward explicit reasoning but may require indirect techniques such as thought experiments and other rhetorical manoeuvres (as discussed above). Secondly, a religious conversion will cause someone to have a different phenomenology (e.g., one sees things as the creations of God, or sees actions in terms of theological good and evil, etc.). Such a change can be sudden and can be striking, even in the scientific context where adopting a new disciplinary matrix, for example Newtonian mechanics in the early eighteenth century, may allow one to see a vast range of phenomena in new (Newtonian) terms. In such cases one may acquire a particularly fruitful, widely applicable set of schemata.

A superficially rather different approach to world-change is the dynamic, neo-Kantian one developed by Paul Hoyningen-Huene (1993). Kuhn (1991, p. 12) described himself as a Kantian with moveable categories. The idea is that in Kant the categories of thought and the nature of the forms of intuition are fixed parts of human nature. Kuhn's proposal is that they are not fixed but changeable and in particular are changeable in response to revolutionary changes in scientific commitments. Corresponding to Kant's phenomena and noumena (things-in-themselves), Hoyningen-Huene draws a distinction between the phenomenal world and the world-in-itself. What changes when there is a Kuhnian world-change is the former, while the world-in-itself stays fixed.

Although I used to think that this was a very different understanding of world-change, I am now inclined to think that this neo-Kantian view and my naturalistic one may be reconciled. If the structure and form of intuition and the categories do not stay fixed, but may change as our paradigms change, we may then ask *how* they change? By what means do paradigm changes affect intuition and the categories of thought? Only psychology can answer that question, and the discussion above touches on the many contributions psychologists can be considered to have made to that problem. As Hoyningen-Huene points out, this combined naturalistic neo-Kantian Kuhn faces a tension in that a properly

[11] Phenomenology has been disadvantaged by its anti-naturalism and it anti-psychologism. Were it to take a greater interest in the psychological mechanisms whereby the world as we experience it is imbued with meaning, it might both add credibility to its claims and at the same time discover a tool for making those claims more precise and rigorous.

conducted psychology, especially if it engages with neuroscience, as Kuhn assumed it must, will be talking about entities and processes that lie more on the side of the world-in-itself than the phenomenal world. Kant and Kuhn were both sceptics about knowledge of things-in-themselves. This is the primary obstacle to developing a combined view. However, if the world-in-itself is just whatever lies beyond the phenomenal world, I do not see that we *must* be sceptical about it. That is, one can make a perfectly good distinction without declaring one half of it to be unknowable (nor the other half to be perfectly knowable). So, maybe, one could develop a naturalistic, neo-Kantianism without built in scepticism vis-à-vis the world-in-itself.

8. CONCLUSION

The proposal in this paper cannot yet claim to be a theory of incommensurability, but may reasonably hope to have indicated what some of the elements of such a theory might be. A fully worked out theory would have to have two phases. First, it would have to characterize carefully the phenomenon of incommensurability. By this I mean we would need to identify certain episodes in the history of science or in communication between contemporary scientists as sharing a set of features, primarily those demonstrating some kind of failure of comparison or of understanding, such that we can reasonably attach the label "instances of incommensurability" to them. Preferably one would do this without hypothesizing about the causes of such cases. Rather the point of this part of the process is principally descriptive. Secondly the theory would elaborate on the various psychological mechanisms mentioned above as possible causes of the instances of incommensurability described in the first phase. Neither phase has been worked out in sufficient detail as yet. While Kuhn and Feyerebend were happy to talk at length about incommensurability, they provided little in the way of detailed historical examination of any instances of it. Furthermore, those cases they do describe are described in terms of their own theories, so it is difficult to tell whether those instances do indeed support those theories or are perhaps better explained by some other theory. There is room also for further work on the psychology of the interaction between contemporary scientists working on a common problem but from different fields. This could yield some valuable data, but the work has not been done as far as I am aware. There is considerable work on the psychology of analogical thinking, but as we have seen, in its application to science Dunbar has focused on the large scale rather than the micro-psychological functioning of analogies in scientific cognitive processes, and in particular we need more data on specifically scientific schemata. So while many of the pieces are available, more work needs to be done on some, and in particular on the interaction between the various pieces. What is significant is that a topic that has been treated in a largely aprioristic, philosophical fashion now needs significant input from a variety of directions and disciplines, from history and sociology of science to various subfields within psychology and cognitive science. The role of philosophy of science is no longer to answer the problem itself but to co-ordinate the contributions of disparate disciplines.

I conclude by noting another distinctive feature of this programme. The incommensurability thesis has been traditionally conceived of as a thesis concerning theories and the languages in which they are couched. For example, Kuhn's (1983, 1991) taxonomic approach to incommensurability identified it with overlapping, non-congruent taxonomies employed by different theories, being his last (published) account of incommensurability following on from a sequence of linguistically and conceptually orientated approaches (Sankey, 1993, 1998). If the psychological account above is along the right lines, then the concentration on theories, concepts, and languages is at least in part misplaced, for it is scientists and scientific communities (or, to be more precise, their cognitive habits) that are incommensurable.[12]

BIBLIOGRAPHY

Andersen, H., Barker, P., and Chen, X. (1996) Kuhn's Mature Philosophy of Science and Cognitive Psychology. *Philosophical Psychology*, 9, 347–364.

Andersen, H., Barker, P., and Chen, X. (2006) *The Cognitive Structure of Scientific Revolutions*. Cambridge: Cambridge University Press.

Anderson, R. C. (1984) The Notion of Schemata and the Educational Enterprise: General Discussion of the Conference. In R. C. Anderson, R. J. Spiro, and W. E. Montague (eds.) *Schooling and the Acquisition of Knowledge*. Hillsdale, NJ: Lawrence Erlbaum.

Bird, A. (2002) Kuhn's Wrong Turning. *Studies in History and Philosophy of Scienc*e, 33, 443–463.

Bird, A. (2005) Naturalizing Kuhn. *Proceedings of the Aristotelian Society*, 105, 109–127.

Carston, R. (2002) *Thoughts and Utterances*. Oxford: Blackwell.

Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1981) Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science*, 5, 121–152.

D'Andrade, R. (1995) *The Development of Cognitive Anthropology*, Cambridge: Cambridge University Press.

Dunbar, K. (1996) How Scientists Really Reason. In R. Sternberg and J. Davidson (eds.) *The Nature of Insight*. Cambridge, MA: MIT, pp. 365–395.

Dunbar, K. (1999) How Scientists Build Models: In Vivo Science as a Window on the Scientific Mind. In L. Magnani, N. J. Nersessian, and P. Thagard (eds.) *Model-Based Reasoning in Scientific Discovery*, New York: Kluwer/Plenum, pp. 85–99.

Forrester, J. (1996) If *p*, Then What? Thinking in Cases. *History of the Human Sciences*, 3, 1–25.

Gentner, D. and Jeziorski, M. (1993) The Shift from Metaphor to Analogy in Western Science. In A. Ortony (ed.) *Metaphor and Thought*, 2nd ed. Cambridge, MA: Cambridge University Press, pp. 447–480.

Gick, M. L. and Holyoak, K. J. (1983) Schema Induction and Analogical Transfer. *Cognitive Psychology*, 15, 1–38.

Hesse, M. B. (1966). *Models and Analogies in Science*. Notre Dame, IN: University of Notre Dame Press.

Holyoak, K. J. and Thagard, P. (1995) *Mental Leaps: Analogy in Creative Thought*. Cambridge, MA: MIT.

Holyoak, K. J. and Thagard, P. (1997) The Analogical Mind. *American Psychologist*, 52, 35–44.

[12] This is not to say that the approach proposed is inconsistent with any elaboration of semantic incommensurability. The discussion of Sect. 5 itself is amenable to development in that direction. In my view Kuhn's philosophical accounts of semantic incommensurability have not been successful. Nonetheless more recently Andersen et al. (2006) (and in various articles since their (1996)) have developed a naturalistic account of semantic incommensurability informed by cognitive science that could be thought to parallel my account of psychological incommensurability. That said the two approaches are logically independent – the present approach aims in part to show that one *can* understand incommensurability without considering issues of language and meaning. I note finally that the central place of the latter in human cognition and psychology does not entail that every significant cognitive phenomenon needs a linguistic explanation; incommensurability may well turn out to be such a case.

Hoyningen-Huene, P. (1993). *Reconstructing Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Hoyningen-Huene, P. and Sankey, H. (eds.) (2001) *Incommensurability and Related Matters*. Dordrecht, The Netherlands: Kluwer.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Kuhn, T. S. (1983) Commensurability, Communicability, Comparability. In P. D. Asquith and T. Nickles (eds.) *PSA 1982*, Vol. 2. East Lansing, MI: Philosophy of Science Association, pp. 669–688.

Kuhn, T. S. (1991) The Road Since Structure. In A. Fine, M. Forbes, and L. Wessels (eds.) *PSA 1990*, Vol. 2. East Lansing, MI: Philosophy of Science Association, pp. 2–13.

Kuhn, T. S. (1992) The Trouble with the Historical Philosophy of Science. *Robert and Maurine Rothschild Distinguished Lecture 19 November 1991. An Occasional Publication of the Department of the History of Science*. Cambridge, MA: Harvard University Press.

Leake, D. (1998) Case-Based Reasoning. In W. Bechtel and G. Graham (eds.) *A Companion to Cognitive Science*. Oxford: Blackwell, pp. 465–476.

Margolis, H. (1987) *Patterns, Thinking, and Cognition. A Theory of Judgment*. Chicago, IL: University of Chicago Press.

Nersessian, N. J. (2003) Kuhn, Conceptual Change, and Cognitive Science. In Nickles (2003b), pp. 179–211.

Nersessian, N. J. (2001) Concept Formation and Commensurability. In Hoyningen-Huene and Sankey (2001), pp. 275–301.

Nickles, T. (2003a) Normal Science: From Logic to Case-Based and Model-Based Reasoning. In Nickles (2003b), pp. 142–177.

Nickles, T. (ed.) (2003b) *Thomas Kuhn*. Cambridge: Cambridge University Press.

Sankey, H. (1993) Kuhn's Changing Concept of Incommensurability. *British Journal for the Philosophy of Science*, 44, 759–774.

Sankey, H. (1994) *The Incommensurability Thesis*. Aldershot: Avebury.

Sankey, H. (1997) Incommensurability: The Current State of Play. *Theoria*, 12, 425–445.

Sankey, H. (1998) Taxonomic Incommensurability. *International Studies in the Philosophy of Science*, 12, 7–16.

Schallert, D. L. (1982) The Significance of Knowledge: A Synthesis of Research Related to Schema Theory. In W. Otto and S. White (eds.) *Reading Expository Material*. New York: Academic.

Sperber, D. and Wilson, D. (1995) *Relevance*, 2nd ed. Oxford: Blackwell.

Stein, N. L. and Trabasso, T. (1982) What's in a Story? An Aapproach to Comprehension. In R. Glaser (ed.) *Advances in the Psychology of Instruction*, Vol. 2, Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 213–268.

COMMENTARY ON BIRD'S PAPER

PAUL HOYNINGEN-HUENE

It is a pleasure to comment on Professor Bird's paper because it is part of the excellent work he has done on Kuhn's philosophy (most importantly Bird, 2000, 2002, 2003, 2004, 2005). His work is especially valuable to me because for the first time an author has provided a fully worked out interpretation of Kuhn that appears to be a real alternative to my own. His alternative interpretation has been intellectually liberating, offering a possible exit sign out of a reading of Kuhn that appeared unavoidable to me. Bird is entirely justified in stressing those elements in Kuhn's theory that he calls "naturalistic", providing a contrast to those elements that I stressed and that may be termed "Neo-Kantian".

However, in the present paper Bird suggests an exciting new possibility that neither he nor I have seriously considered before, namely a possible reconciliation between the naturalistic and the Neo-Kantian view of Kuhn regarding world change (p. 36). In this commentary, I would like to pursue this line because it appears to be fruitful. Let me begin by first pinning down the contrast that is at issue.

Bird's approach is naturalistic. This means that he takes up, develops and amends those elements that Kuhn takes from the sciences, especially from psychology and the neurosciences, and uses at the meta-level, i.e., in order to analyse the (natural) sciences. Bird's main interest in the present paper is to develop incommensurability. More to the point, he is interested in a naturalist, i.e., empirical analysis of those cognitive capacities and structures that scientists use when thinking, and that may give rise to incommensurability in particular circumstances. His result is that certain quasi-intuitive cognitive capacities (QICC's) that a scientist has acquired due to social induction in a particular field are paradigm dependent. Due to this dependence, various phenomena known from the debate about incommensurability are then expected to occur across the revolutionary divide – "a certain kind of incommensurability" prevails (p. 26). When contrasted with an aprioristic approach, the specifics of this naturalistic approach to incommensurability are:

- The relevant features are not a priori characteristics of scientific cognition but must be empirically investigated.
- The relevant features "exist in virtue of the nature of human psychology" (p. 27). This does not preclude Incommensurability being a social phenomenon because scientists belonging to the same tradition have all received an education that is similar in decisive respects.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 41–46.

What seems to be at issue here is roughly the following. The two parties, let's call them "naturalists" and "apioricists", agree that there are features of the human mind that give rise to the phenomenon of incommensurability. They may disagree about some details of incommensurability, but this may be neglected here. The main difference between the two parties concerns how these features of the human mind can and should be investigated. The difference between the two parties thus appears at first sight epistemological and not ontological.[1] The naturalists claim that only an empirical investigation can disclose the existence and the nature of those features. In principle, the aprioricists might claim that an empirical investigation of those features is either impossible (because they are only reflectively accessible), or not necessary (because there is a better way to access them), or misleading (because an empirical investigation misrepresents the nature of those features), or useless (because their intended function depends on their being determined a priori). Let us look at these possibilities in turn.

For the sake of argument, let us assume that the aprioricists have identified, in an a priori way, certain features of the mind that give rise to incommensurability under appropriate conditions. First, is it plausible to assume that the naturalists can have no cognitive access whatsoever to these features? However deep down in the human mind these features may reside, they must have some observable effects, and it is a reflective analysis of these effects that leads the apioricists to these features, in the very best case in a secure, a priori way. However, the naturalist may be unaware of this a priori possibility, or he may distrust it. Instead, on the basis of the observed effects, he will form an empirical hypothesis about the features of the mind that are responsible for these effects. Of course, this hypothesis can be a postulate of the existence of just those features of the mind that the aprioricist has identified in an a priori way. Therefore, it is implausible that the naturalist cannot have cognitive access to features of the mind that are knowable a priori.

Second, is an empirical investigation of features of the mind that are knowable a priori really unnecessary as an aprioricist may claim? Of course, the aprioricist is right here. For instance, an empirical investigation of the sum of the angles of planar triangles is indeed unnecessary because we have an absolutely watertight proof for its being 180°. (However, should the slightest doubt creep into the supposed solidity of our a priori reasoning, an additional empirical investigation may prove helpful.) In defence of the naturalist, we may state that an empirical investigation of features of the mind that are knowable a priori, is not harmful in itself – it is just unnecessary.

Third, is an empirical investigation of features of the mind that are knowable a priori misleading because it misrepresents the nature of these features? The result of the a priori investigation of those features is that they are necessary features of the mind, perhaps constitutive for what a mind is, whereas the empirical investigation resulting in the same features cannot state their necessity. Of course, it does make a difference whether, for example, the equality of inertial mass and gravitational mass

[1] I am saying "*at first sight* epistemological" because the different epistemological approaches are founded in differences assigned to the objects of the approach, i.e., the pertinent features of the human mind. For the naturalist, these features of the human mind are contingent properties of the mind (or, at least, methodologically they should be treated as such) whereas for the aprioricist, they are necessary (or essential) properties of the mind. On the relevance of this difference, see the third remark below.

has been experimentally shown or demonstrated as a necessity. For instance, if the latter is (really!) true, no further experiments testing the equality with higher accuracy will be conducted. However, in many other situations the equality will simply be taken as a fact, irrespectively of its epistemological status. Similarly, the epistemological properties of features of the mind that give rise to incommensurability will be largely inconsequential, given that they are really there. In the first case, incommensurability would be a necessary consequence of the operations of the human mind in certain situations – it could not be otherwise. In the second case, incommensurability would be a contingent consequence of the operations of the human mind in certain situations – it could be otherwise, if the mind were different.

Fourth, is an empirical investigation of features of the mind that are knowable a priori useless because the intended function of these features depends on their being determined a priori? This may indeed be the case. For instance, in an ambitious foundationalist epistemological project that is intended to be non-circular, it may be necessary to determine certain features of the mind a priori. This holds for features of the mind that serve to found the possibility of empirical knowledge. In the context of such an aprioricist project (like Kant's), an empirical determination of those features in principle misses the intended goal of the project. If empirical knowledge should be put on firm grounds because it is dubitable, then this ground must clearly not contain any of this dubitable empirical knowledge. However, in the context of less strict epistemological projects, empirical knowledge of features of the mind may indeed be admitted. At least the early Kuhn was certainly not involved in a strictly foundationalist, non-circular epistemological project as his liberal use of empirical knowledge in building up and making plausible his theory demonstrates. The tendency of the later Kuhn to prefer a priori principles to empirical knowledge seems less due to increased methodological strictness of his project as due to an intended gain of independence from historical data and their fluctuating interpretations. Be that as it may, it is clear that in some ambitious epistemological projects knowledge of some features of the mind must be a priori. In other less ambitious projects, it need not be. Of course, the proponents of those ambitious, non-circular, foundationalist projects have the burden of demonstrating the feasibility of their task, especially regarding the content of their a priori claims about features of the human mind. Because the difficulties of such projects are so colossal, many philosophers have followed more modest lines. However, in Kuhn's case, the price to be paid was a substantial lack of clarity regarding the epistemological status of his theory (see Hoyningen-Huene, 1993, Sect. 3.8).

The result of this discussion is that the naturalist and the aprioricist can enjoy a sort of peaceful competition, unless the aprioricist pursues a strictly foundationalist, non-circular epistemological project. To the naturalist, all valid results gained by the aprioricist about features of the mind are welcome because there are no obstacles to accepting a priori truths about the mind, and they free him from some empirical work. To the aprioricist, all valid results gained by the naturalist are heuristically useful because empirically valid results about the mind challenge him to find a priori justifications for them. Regarding at least some aspects of incommensurability, the naturalist and the aprioricist can fruitfully cooperate because nothing hinges on the epistemological status of those features of the mind that give rise to incommensurability. Whether incommensurability is a contingent or a necessary consequence of the constitution of the human mind is inconsequential at least in those cases in which the presence of incommensurability is uncontroversial.

So, indeed, at least a partial reconciliation between the naturalist and the aprioricist appears to be possible regarding incommensurability. However, this is not the whole story. As is indicated by the labels that Bird uses to characterize the two pertinent positions, namely "neo-Kantianism" and "naturalism", more is at issue. We are not just dealing with an epistemological issue about the appropriate approach to those features of the human mind that are responsible for incommensurability. Rather, ontological questions come into play. At least historically, the introduction of incommensurability by Kuhn came with a challenge to realism. Kantian positions challenge ordinary realism by claiming the presence of genetically subject-sided factors in what we call real – not what we call real by mistake or error, but what is *really* real to us, like mountains, coffee or books. Kant determined these genetically subject-sided contributions to the real as time independent and as constitutive of thinghood (what is common to *all* things), and he identified them as the so-called forms of intuition, namely time and space, and his categories. Neo-Kantians vary Kant's original idea in one or another way. On the neo-Kantian reading, Kuhn uses historically variable similarity and dissimilarity relations that are not constitutive of thinghood itself, but of particular things, of specific classes of things, and of specific classes of situations.[2] At any rate, all Kantian positions identify the real with a phenomenal world, i.e., a world whose objects contain genetically subject-sided factors. And all Kantian positions deny, in one way or another, that we are able to subtract the genetically subject-sided contributions from the real and look at what is left (if there is anything left). Realists, on the other hand, deny that genetically subject-sided contributions are so deeply immersed into what appears to be the real that they cannot be theoretically subtracted. In their view, it is one of the main functions of scientific theories to subtract genetically subject-sided contributions from phenomena (like colours) to arrive at the real (like electromagnetic waves with a certain distribution of wavelengths). According to the realist, successful scientific theories lead

[2] It should be noted at this point that from a strictly Kantian point of view, Kuhn's approach is neither really Kantian nor can it be the last word. Kuhn applies his similarity and dissimilarity relations to aspects of things, things, and constellations of things, and at least the aspects of things must somehow be there and accessible to the epistemic subject before the similarity and dissimilarity relations can be applied. Kuhn treats them as if they are simply given to the epistemic subject from outside, and he treats time, space and causality in the same way (see Hoyningen-Huene, 1993, Sects. 2.1a and 3.2). This is, of course, unacceptable from a strictly Kantian point of view. In his approach, Kuhn does not even touch the main Kantian theme, namely the constitution of thinghood by means of genetically subject-sided elements (the forms of intuition and Kant's categories). Kuhn does not analyze the constitution of thinghood in terms of genetically subject-sided elements as Kant does, but asks, *given* thinghood, how are different things, different classes of things, and different classes of situations constituted by means of genetically subject-sided elements? It is thus somewhat misleading to write with respect to Kuhn, as Bird does on p. 36: "If the structure and the form of intuition and the categories do not stay fixed, but may change as our paradigms change, […]". In a strictly Kantian view, Kuhn analyzes the effects of genetically subject-sided elements that come into play *after* Kant's forms of intuitions and his categories have done their work whose result is epistemically accessible thinghood. (I have discussed the dependence of Kuhn's theory upon the preceding availability of thinghood [wherever it may come from] and upon other presuppositions in Hoyningen-Huene (1993, pp. 78–81) – For the sake of clarity, let me stress that I am not endorsing Kant's position at this point, but that I am only relating it to Kuhn's position.

us, albeit in a fallible way, from the phenomena to the purely object-sided, i.e., to the real in an absolute (and ordinary) sense, at least approximately.

The contrast made to neo-Kantianism (and to Kant's original position) that I just invoked is *realism*, not naturalism, as compared to Bird's contrast. It appears that naturalism is not the proper contrast to neo-Kantianism, that it is somehow askew to oppose these positions. However, this impression may seem to dissolve once one realizes that most scientists are indeed realists in that they never doubt, at least not in their professional life, the absolute reality of what they investigate. Of course, falling bodies, water, animals, tectonic plates, the Moon, the brain, and so on are typically conceptualized in science as being completely independent of our theorizing, and thus as real in an absolute sense. This realist attitude is part and parcel of what Bird means by "naturalism", even if this is not entirely explicit in the present paper (but see, e.g., Bird, 2000, pp. 210–211, 266). However, one should remember that some natural scientists, among them especially some quantum physicists, are non-realists in the sense that they believe that even some of our most basic categories like thing, property of a thing, etc. are not just out there (i.e., are purely object-sided) but are of our making (i.e., have genetically subject-sided contributions).[3] It is important to note that for these physicists, non-realism is not a contingent philosophical addition to their science, but rather a necessary consequence of it. I am aware of the fact that these matters are highly controversial but at least they show that realism is not a view that is conceptually and therefore necessarily linked to the scientific attitude. Thus we should not treat realism as a necessary ingredient of naturalism.

So the real issue is not naturalism versus Neo-Kantianism, because non-foundationalist forms of Neo-Kantianism are quite compatible with naturalism. The real issue emerging from the incommensurability debate is the alternative between realism and non-realism (of some Neo-Kantian flavour). Because non-realism is obviously compatible with incommensurability (although one may dislike non-realism for other reasons, among them perhaps its implausibility to common sense), it seems that the following question is the crucial one. Assuming the phenomenon of incommensurability in the history of science, and given that incommensurability contains world change in some plausible, non-superficial interpretation (e.g., something like Bird's), is then some form of realism a plausible option? I have the impression that most realists do not ask this question, but approach the problem differently. Their attitude is this: Given my form of realism, how must I interpret incommensurability and its concomitant world change such that they either disappear as merely metaphorical or superficial, or that they come out as somehow real but that they don't threaten realism. Of course, in this way, realism will come away unscathed – but only because it has not been challenged. However, philosophically this is not an interesting result. Here is a different attitude that may prove to be more fruitful. Let us investigate incommensurability and world change in a non-presentist way, i.e., how they are experienced by acting scientists in their historical settings. Let us use all scientific (or philosophical, if available) means that help us in this investigation, especially psychology. Regarding the analysis

[3] Einstein is also a case in point, although not for quantum-theoretical reasons. See Einstein (1949, esp. pp. 673–674) and Rosenthal-Schneider (1949) for commentary and further references.

of world change, let us try *not* to presuppose philosophical positions like realism or non-realism. Instead, let us reconstruct what the world views are before and after the revolution. After consideration of several such cases of incommensurability, let us ask the following question: Given that our epistemic position as analysts of scientific change is not fundamentally different from the epistemic position of the scientific subjects investigated, are these cases of world change compatible with a plausible realist position? Speaking for myself, I may state that at least my confidence in any form of realism is seriously undermined by this experience, and it is my impression that many competent historians feel the same. Of course, these are not decisive arguments, but this way of looking at the situation may open up a discussion in which many seem to be too sure of their own position.

Acknowledgements

I wish to thank Eric Oberheim for bringing the references in note 3 to my attention and for stylistic improvements.

BIBLIOGRAPHY

Bird, A. (2000) *Thomas Kuhn*. Princeton, NJ: Princeton University Press.

Bird, A. (2002) Kuhn's Wrong Turning. *Studies in History and Philosophy of Science*, 33, 443–463.

Bird, A. (2003) Kuhn, Nominalism, and Empiricism. *Philosophy of Science*, 70(4), 690–719.

Bird, A. (2004) Kuhn, Naturalism, and the Positivist Legacy. *Studies in History and Philosophy of Science*, 35, 337–356.

Bird, A. (2005) Naturalizing Kuhn. *Proceedings of the Aristotelian Society*, 105, 109–127.

Einstein, A. (1949) Remarks Concerning the Essays Brought Together in this Co-operative Volume. In: P. A. Schilpp (ed.). *Albert Einstein: Philosopher-Scientist,* Vol. 2, La Salle: Open Court, pp. 665–688.

Hoyningen-Huene, P. (1993) *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. Chicago, IL: University of Chicago Press.

Rosenthal-Schneider, I. (1949) Presuppositions and Anticipations in Einstein's Physics. In: P. A. Schilpp (ed.). *Albert Einstein: Philosopher-Scientist*, Vol. 1, La Salle: Open Court, pp. 131–146.

PART 2

INCOMMENSURABILITY IN A WITTGENSTEINIAN PERSPECTIVE: HOW TO MAKE SENSE OF NONSENSE

NONSENSE AND PARADIGM CHANGE*

ARISTIDES BALTAS

Abstract This paper attempts to show how Kuhn's and Wittgenstein's works can be of mutual assistance despite their apparent heterogeneity. One face of this project is to analyse the conceptual aspects of paradigm change in physics as described by Kuhn with the help of Wittgensteinian tools, especially what he calls "nonsense" and "grammar", as well as the metaphors of the "hinges" and the "ladder" he employs. On this basis, the paper investigates the process through which a teacher can teach radically new concepts to science students still entrenched in the old incommensurable scientific paradigm. It articulates the thesis according to which the judicious use of nonsense is the indispensable ladder for the elucidation of the novel concepts and for the acquisition of the novel paradigm. The other face of the project is, reciprocally, to use Kuhnian concepts in order to attain a better understanding of Wittgenstein's undertaking. From this side, the paper attempts showing that Wittgenstein's early philosophy in the *Tractatus* can be understood as a major (perhaps the most) radical "paradigm shift" for the whole "disciplinary matrix" of philosophy. This paradigm change in philosophy is of a very particular kind since it aims to silence philosophy in its entirety by erasing all philosophical problems without exception.

Keywords Wittgenstein, Tractatus Logico-Philosophicus, nonsense, grammar, paradigm change.

1. INTRODUCTION

The present paper has three parts. In the first part we will try arguing that the judicious use of nonsense is indispensable for elucidating the fundamental concepts of a radically novel scientific paradigm to students still entrenched in the superseded one. This, we will take it, is the only way available for making the students enter the novel paradigm and come to understand its workings. To bring out the salient points as simply as possible, we will focus on only one example, namely the passage from classical mechanics (CM) to the special theory of relativity (STR). In order not to burden the reader with too much physics, we will simplify drastically the issue and take the passage in question

* It is a pleasure to thank Jim Conant and Spyros Petrounakos for reading the present paper and commenting generously on it. Some of their remarks suggest important new avenues for developing the ideas presented here; however it is impossible to try walking on these avenues at this moment.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 49–70.

as fundamentally concerning the understanding of just one novel concept, that of the electromagnetic field. The discussion will be based on work done in Baltas (2004), where I try bringing out the grammatical aspects involved in radical paradigm change by basing myself on the conception of grammar developed in Wittgenstein's later work. A summary of the ideas put forth in that paper will be provided to start with.

The relation between elucidation, nonsense and understanding we have in mind here derives from Wittgenstein's *Tractatus Logico-Philosophicus* (Witgenstein 1986, from now on TLP or simply Tractatus) which, however, appears as totally indifferent to paradigm change and to all the related issues. In the second part of the paper then, we will try arguing that, such appearances notwithstanding, this early work of Wittgenstein's does countenance radical paradigm change in science but that it countenances it in its own very particular way: the conceptual and grammatical issues involved are brought down to the underlying logic thereby losing much of their philosophical bite.

This is an acknowledgement of a philosophical issue that is also a dismissal, an acknowledgement *cum* dismissal which, according to some commentators,[1] characterizes the way in which the *Tractatus* countenances all philosophical issues. This is to say that the *Tractatus* presumes putting to rest, after acknowledging them, all major philosophical problems without exception while erasing fully at the same time all philosophical theses which the *Tractatus* itself appears as forwarding. This is a way of putting an end to philosophy altogether and, accordingly, the *Tractatus* presents itself as constituting a major, perhaps the most radical, "paradigm change" not in science but in philosophy.

The quotes in "paradigm change" intend underlining the fact that science and philosophy form very different endeavors and hence that the results attained from a philosophical study of science – as well as the vocabulary helping us to attain such results – cannot be applied *mutatis mutandis* to our understanding of philosophy; at best, such a study can only hope of coming up with an elucidatory analogy. Be that as it may however, and while trying to take into account all relevant differences, we will argue in the third part of the paper that the *Tractatus* can profitably be viewed as indeed attempting the most radical "paradigm change" in the whole "disciplinary matrix" of philosophy, the "paradigm change" aiming no less than to silence philosophy in its entirety.

Such a reading of the *Tractatus* gives preeminence to the famous penultimate paragraph of that work: "My propositions are elucidatory in this way: he who understands me finally recognizes them as nonsensical, when he has climbed out through them, on them, over them. (He must, so to speak, throw away the ladder, after he has climbed up on it.) He must surmount these propositions; then he sees the world rightly" (TLP 6.54). In the third part of the paper then, we will try exploiting the results of the first two in order to throw some light on that paragraph. The kind of questions we will be asking will be: how exactly does the relation between elucidation and nonsense, as we will have been discussing it for the case of physics, can help us understand what Wittgenstein means in the paragraph in question? More generally, how can the results of the first two parts help us understand early Wittgenstein's conception of his own work and of philosophy in general? How can our author write a treatise that invites its reader at the end to recognize all the propositions of that treatise as nonsensical? What is the purpose for writing such a treatise in the first place?

[1] See the discussion in Crary and Read (2000) as well as Ostrow (2002).

The aim of that last part of the paper is not only our trying to understand what the *Tractatus* per se is all about. We will conclude the paper by saying a few words as to how such understanding can shed new light on the ways in which Kuhn's work in philosophy of science can relate to what the *Tractatus* has to say about philosophy in general and hence how these two apparently so disparate works can be of mutual assistance.

2. PART I: NONSENSE AND PARADIGM CHANGE IN PHYSICS

Put as simply as possible, the physics involved in the passage from CM to STR is more or less the following.

Maxwell's electromagnetic theory (1864) achieved, among other things, the unification of electricity, magnetism and optics. One of the more salient results of that achievement was the prediction of electromagnetic waves, the existence of which was later confirmed experimentally by Hertz. What we should note for our purposes is that, for the wisdom of the period, these waves, by being precisely waves – the propagation of a medium's disturbances according to the classical definition – necessarily required the existence of the corresponding material carrier. Honoring a long tradition, this carrier was then called the ether. However, all efforts to pin down the ether's properties failed for no obvious reason, its character thereby becoming more and more elusive and mysterious. A crisis in Kuhn's (1962) sense of the term thus settled in with no obvious way out. It was in this context that in 1905 Einstein published his "On the Electromagnetics of Moving Bodies", whereby he introduced STR.[2] Stating it as simply as possible, no ether was required according to STR, for the electromagnetic waves manifested the existence of a new kind of physical entity, the electromagnetic *field*, an entity altogether different from particles and media. The perhaps most prominent characteristic of the electromagnetic field is precisely the fact that it can move by itself in vacuum without the support of any medium whatsoever. In Einstein's word, the ether thus proves "superfluous". Accordingly, understanding STR hinges on understanding the electromagnetic field, given the fact that such an entity is entirely inconceivable for CM.

Why is it inconceivable? Simply because the electromagnetic waves that manifest the field's existence continue to be waves and a wave is classically *defined*, as we have just said, as the propagation of a medium's disturbances. This definition is not the result of a voluntary free decision that gives content, as it were, to a newly introduced technical term. This is a definition which intends to capture – and does indeed capture – all our everyday experience with ordinary waves (ocean waves, sound waves, etc.) as well as all what the practice of (classical) physics had come up with in that respect up to advent of STR. In other words "wave" merely names, with all the necessary mathematical niceties, a particular kind of movement that is followed by the disturbances of a medium and hence the physical concept "wave" is *analytically* related to the physical

[2] Holton (1988) presents compelling evidence to the effect that *c*.1905 Einstein was working in relative isolation from the scientific community. For this reason, it is more or less historically inaccurate to consider his 1905 relativity paper as a self-conscious intervention to what we have just called a crisis situation in physics. The paper can be read as that only post festum and only from a philosophical perspective which is not concerned overmuch with historical details.

concept "medium". It follows that the alleged existence of waves that can propagate in vacuum, in absence of the material medium *whose disturbance they nevertheless are by definition*, amounts to a quintessential contradiction. It is like maintaining that some bachelors may be married males after having *defined* bachelor precisely as an unmarried male.

3. ON THE GRAMMAR OF PARADIGM CHANGE

In Baltas (2004) I have tried to present a grammatical account analyzing how contradictions like this arise within history of science and how overcoming them – what amounts to the conceptual aspect of radical scientific discovery – boils down to the introduction of a novel, previously inconceivable, *conceptual distinction* and to the concomitant *widening of the grammatical space available to the inquiry*. Within this wider grammatical space the contradiction disappears: the breakthrough of paradigm change arrives to establish a logically consistent and fully grammatical conceptual system centered on the newly introduced conceptual distinction. This is a novel conceptual system, radically different from the old.

The main idea is the following. Any scientific conceptual system of the sort interesting us here (like that of CM) harbors background "assumptions" which can be pictured as the grammatical "hinges" (Wittgenstein, 1969, paragraphs 341–343) that silently fasten from the background our understanding of it. A crisis situation settles in when the process of inquiry stumbles unawares on some such "assumptions". Such "stumbling" induces a crisis situation because the "assumptions" at issue are being entertained completely unawares in the sense that, even if somebody arrives to state them, they are taken as totally obvious and self-evident and hence invulnerable to their being meaningfully questioned; it is these that assure the overall grammatical coherence of the old conceptual system. In other words, the *grammatical possibility* of their being questioned doesn't arise in normal[3] circumstances while *resistance* to its arising is also at work.[4]

As Wittgenstein makes clear in *On Certainty*, the existence and silent functioning of such "assumptions" is essential and unavoidable. In that work, however, what is mainly emphasized is their *positive* role as, precisely, the indispensable grammatical "hinges" that have to stay put for the door of understanding to move. The additional point of our account is that *the very same* background "assumptions", *the very same* grammatical "hinges" are not only *necessary for understanding* but, on occasions, constitute as well the fundamental *obstacles to understanding*. Hence, to keep the same metaphor, the breakthrough of paradigm change amounts to our realizing that some of these "hinges" have turned rusty. Our realizing this goes together with the introduction of a novel conceptual distinction, inconceivable within the grammatical space determined by

[3] "Normal" is here intended to bear the connotations of "normal science" in the sense of Kuhn (1962).

[4] In the case at hand, these are precisely the background 'assumptions' fastening the classical definition of "wave". Resistance to their being questioned is encountered because those still entrenched in CM do not see that *there can be something "beyond" classical physics* that would change drastically some of their most fundamental conceptions and therefore they perceive their interlocutor as trying *per impossibile* to tear apart the analytic relation defining "wave" through "material medium".

the old "hinges" (in the case at hand the distinction between "classical" waves and "electromagnetic waves" which *can* propagate in absence of any material carrier) and the establishment of a novel conceptual system (that of STR) centered on this distinction. Advancing this conceptual distinction and exploring its consequences is tantamount to removing the old "hinges" implicated and simultaneously replacing them by others, those fastening our understanding of the breakthrough that is simultaneously taking place. The new concepts (here: wave, time, space, mass, energy, etc.) appear as nonsensical from the vantage point of the superseded paradigm for their meaning, as determined by the novel "hinges", is radically at odds with the deeply entrenched intuitions assuring the meaning of the old concepts on the basis of the old "hinges".

Removing the old "hinges" is equivalent to widening the grammatical space available to the inquiry. This wider grammatical space can host a *re*interpretation of the old conceptual system as well as, in some cases like the one at hand, an imperfect rendition of parts of the old conceptual system in terms of the new. Hence, within the bounds set by the novel conceptual system, these parts – precisely as reinterpreted – can sometimes still remain useful to scientific practice.[5] However, the reverse movement is blocked, for the old, narrower, grammatical space cannot accommodate what the new allows. This makes the two succeeding paradigms asymmetrical while the incommensurability between the two, in Kuhn's (1962) sense of the term, is based on this asymmetry. The phenomena of communication breakdown discussed in the literature are due precisely to this asymmetry: those not having gone through such radical conceptual reshuffling cannot but still take blindly for granted the background "assumptions" fastening the understanding of the old paradigm.

After the conceptual tumult of paradigm change has settled down and the new conceptual system has become established, the wider grammatical space subtending it has become instituted *for good*: on the one hand, all further inquiry is based on it while, on the other, returning to the old, narrower, grammatical space becomes impossible. For it is obviously impossible to push back into the background the novel grammatical possibility that has become disclosed, the possibility that has allowed us not only to come to terms with the crisis situation but also and more significantly to promote successful research in the directions it has opened. Such widening of the grammatical space is therefore an *irreversible* achievement and hence theory choice and all the attendant issues can be negotiated only on the basis of the novel grammatical space.

After this wider grammatical space has proved again and again its capacity to host novel, previously inconceivable, scientific results and after the attendant grammatical peace has become instituted, practicing scientists tend to look condescendingly upon the replaced paradigm: the old contradiction becomes retrospectively interpreted as a kind of oversight whose gripping power had been manifesting only *our* failure of noticing a grammatical possibility that could not but exist to begin with. This is our alleged failing of having come up with the conceptual distinction at issue well before we actually did and this is the shortcoming of our scientific past in respect to the present. The success of the breakthrough in question is taken to demonstrate, that is, that the this grammatical

[5] These are the cases where, in the idiom of physicists, the relevant parts of the novel conceptual system possess a "classical limit".

possibility could not but have been always already lying "out there", waiting, precisely, to be *dis-covered* by the cleverest of us. The quasi inevitable whiggism of practicing scientists finds here its root as well as some apparent reasons for its justification.

It should be clear that this is a wrong appraisal of the relation between scientific present and scientific past and that no cleverness whatsoever can be involved here. Coming to see a novel grammatical possibility in the way that Einstein, for example, did cannot but be an affair of allowing our imagination free rein; but within the scientific activity, imagination can neither be appealed to nor be relied upon unless it becomes backed up by particularly solid scientific success, the kind of incontrovertible success required by the stringent criteria of this particularly sober and highly disciplined activity. Imaginative proposals of all kinds may well be put forth in a crisis situation but only very few of them can be justified ex post facto by rock-hard results that force a widening of the grammatical space in the ways we have been discussing. The others remain mere flights of the imagination, perhaps admirable in themselves, but which, insofar as they don't bear scientific effects, are lost to science and soon forgotten.

On the other hand, grammatical possibilities cannot just lie "out there". Grammar in the sense intended here forms the "riverbed" of language (Wittgenstein 1969, Paragraphs, 97 ff.), "dictating" from the background the latitude of its meaningful uses. This is to say that a novel grammatical possibility does not exist "in itself", passively waiting its discovery. The "riverbed" can change and a new grammatical possibility can come about only when some of our practices (which are never merely linguistic) arrive to impose it in ways which would need a careful study in their own right.

We should add that effectively widening the grammatical space and hence regaining grammatical coherence in the ways we have been describing is at one and the same time reinstating *logical* consistency: coming up with grammatical room for the novel conceptual distinction is concurrently eliminating the logical contradiction. Putnam (2000) discusses a simple Wittgensteinian example which, although concerning only colloquial language in its everyday uses, may be suggestive of what happens in paradigm change. Thus the logical standoff forced by the contradictory order "Come to the ball both wearing and not wearing clothes!" can be overpowered by the addressee's imagining the possibility of her wearing a fishnet. The grammatical space subtending the linguistic fragment "wearing clothes" within our everyday practices becomes thus enlarged to accommodate a novel distinction, inconceivable except by an act of imagination: before this act arrives to enlarge the set of grammatical possibilities in the context at issue, the "hinges" governing from the background the uses of that linguistic fragment had been making us take silently for granted that wearing or not wearing clothes has nothing to do with fishnets. By the same token, the logical contradiction disappears as such. Of course, what distinguishes scientific practice in this respect is the requirement that the conceptual as well as the experimental implications of such an act of imagination should be born out not just by colloquial language usage (which, as we have implied above, *does not* bear out novel scientific concepts in many cases) but mainly by the workings of the world.

It is important to note for what follows that this widening up of the grammatical space does not affect much what happens outside the scientific process of inquiry itself. Colloquial language does not reform grammatically in such cases while colloquial understanding continues to go on more or less just as before: the novel paradigm remains highly counterintuitive, requiring a long process of instruction, if not initiation, in order to become properly understood. This is to say that, to achieve such understanding, prospective students have to pass through the same kind of experience that scientists have undergone. It should be underlined that this situation is not proper to STR and to modern science in general; pedagogical research has amply demonstrated that even physics graduates continue to conceive spontaneously the world in pre-Galilean terms. Therefore, to teach science effectively, one has to take careful account of such unavoidable resistance while popularized presentations of science tend to flood the market precisely in order to appease it.

4. TEACHER AND STUDENT

Be that as it may, we are obliged to acknowledge that physicists, for their part, have the experience and the tools necessary for coping with situations like this. And hence the question becomes: how can they, how can we start teaching the novel paradigm of STR to someone, to our student, who is still entrenched in the old? How can we make her conceive the electromagnetic field, i.e., the disturbances of nothing that propagate by themselves within nothing? What kind of helpful picture can we draw that could depict such an impossible state of affairs?

To answer these questions we have to note first of all that it would be impossible to make our student understand us through our simply expounding STR and the evidence for it. Since her system of concepts (that of CM as it hooks onto the world) is fundamentally self-consistent while the grammatical space subtending it is narrower than that subtending ours (that of STR), we cannot straightforwardly argue and prove our case the way we could if we were sharing strictly the same grammatical background. The fact that we have entered the conceptual system of STR and understood it has changed the grammatical "hinges" on the basis of which we understand all the relevant concepts, both those of CM and those of STR, while our student has not yet performed the passage. In other words, our student has not yet understood STR and hence she has not yet realized what talking from the vantage point of STR might mean as well as what this might imply or entail. Such talk appears to her as mere nonsense. All this is to say that proof, evidence and all that are *insufficient*, that is per se *powerless*, for doing the job, and this wholly irrespective of anyone's reasoning capacities. It is, we take it, amply demonstrated by our everyday discursive practices that when proof and evidence is thus powerless we resort to, and cannot but resort to, elucidations.[6]

Second, it should be clear that to get our student to understand us is to make her *rationally* forego her conceptual system and join *rationally* ours. No argument from authority and no other kind of psychological force should come into play as means of persuasion. To reach intellectually our student then, the kind of elucidations that we have to forward should be such as to build a kind of bridge, or, if we like, a kind of

[6] This is well argued for by Weiner (1990) where she examines Frege's ways of proceeding in order to make someone not initially familiar with his *Begriffschrift* enter it and, so to speak, understand it from the inside. See also Conant (1992).

"ladder", between the two of us, a bridge or "ladder" that would make the necessary concessions to her not understanding us, a bridge or "ladder", that is, which would enable us to *meet her halfway*.[7] This is a bridge or "ladder" which, to be precisely a means of connecting us two, has to share elements from both systems, employ terms, locutions and pictures which are somehow common to both. Since each system is self-consistent while the two are asymmetrical in the way we discussed above (the grammatical space subtending STR can accommodate an interpretation of the conceptual system of CM while the grammatical space subtending CM cannot accommodate any interpretation of the concepts of STR; on its basis those concepts appear as nonsensical), the bridge or "ladder" connecting the two *cannot but constitute nonsense* from the point of view of either. Elucidation and nonsense thus seem to go hand in hand.

Third, to meet us halfway in the bridge or "ladder" we are trying to build for our student's benefit presupposes a certain kind of "*good will*" (TLP 6.43) on her part. Nobody can teach anything at all to somebody displaying "*bad will*" in the relevant respect, i.e., to somebody unwilling to learn. If obviously we cannot require of our student to "have already thought the thoughts" (TLP Pr., Paragraph 1) that we are trying to teach her – for then she would have no need of us, having paragraph already understood STR – we are compelled nevertheless to require of her to have some inkling toward "similar thoughts" (ibid.), even if only in the form or her activating the "good will" to learn what we are trying to teach her. And this not merely because this is the only way of her responding to the "good will" *we* are laying out in our effort to teach her, but mainly because this is a sine qua non prerequisite of all teaching.

Given all this, let us focus on what can be taken as a perfect example for the kind of answers we are seeking to our questions above. Addressing himself to laymen, that is to prospective students like the one confronting us here, Stephen Weinberg, Nobel laureate in physics and grand master of exposition, "defines" the electromagnetic field as follows: "A field is a taut membrane without the membrane" (Weinberg, 1977).

We maintain that this obviously self-destroying "definition" provides exactly the kind of bridge or "ladder" we are after: its first part makes reference to CM (taut membranes are legitimate objects of that theory) while the second (that the membrane in question, like the ether, does not exist) sits squarely within STR. The fact that the "definition" as a whole destroys itself points at the grammatical asymmetry and hence at the incommensurability gap between the two conceptual systems while the fact that it manages to create some kind of "picture" does elucidate somehow the concept at issue, even if this picture itself is impossible: as the vanishing smile of the Cheshire cat can show (this, we recall, is a smile without a cat),[8] we do have "something" in mind when trying to picture an elusive taut membrane that does not exist. Obviously, when considered in itself, i.e., in no relation *to the purpose for formulating it*, the fact that the "definition" is self-destroying makes it fully nonsensical: it *precludes* by what it is saying what it *presupposes* in order to convey what it says.

7 I owe the idea behind this expression to Diamond (1995b) and to a discussion with Andreas Karitzis. Diamond makes clear that no precise rules can be forthcoming in cases like this and it is necessary that imagination should come into play, in ways similar to its entering paradigm change the ways we have been discussing.

8 I am referring, of course, to Lewis Carroll's *Alice in Wonderland*.

Weinberg's "definition" thus amounts to his constructing one fundamental rung of the "ladder" we are talking about. This is, as it should be, an elucidatory rung *despite* its being nonsensical; or, rather, it is elucidatory *because* it is nonsensical. Its construction thus amounts to Weinberg's *employing nonsense judiciously*, which means employing it *both* in conformity with his purpose *and* with the efficiency required for striking decisively what he is targeting. In one word, Weinberg's nonsensical "definition" arrives indeed to "hit the nail on the head" (TLP Pr., Sect. 7).

Let us try becoming clearer on this. The fact that Weinberg's "definition" destroys itself amounts to its laying bare that it is impossible to understand STR in the terms of CM, for it self-destructs by referring precisely to them. However it is this self-admitted failure that makes the "definition" *effective* in the given context: only by openly exhibiting or "*showing*"[9] that it is impossible to come to understand the STR concept of the electromagnetic field in terms of the classical concepts can the "definition" pave the way for properly understanding the new concept. Hence, to make its point, the "definition" *should* destroy itself and should destroy itself *thoroughly*, to the point of becoming fully nonsensical. In this way and only in this way its very failure becomes the condition for its success. To say the same thing differently, by its very utterance in the given context Weinberg's "definition" *performs in itself and by itself* the particular kind of *not understanding* which our student should submit herself to if she is to understand the new concept. It is such acknowledged submission to *not* understanding (the new concepts in terms of the old) that forms the first step towards proper understanding of the new concept and thereby of the workings of the novel paradigm in its entirety. Such submission amounts to our student's forcing down the *resistance* inevitably arising in her when she hears the nonsense we are advancing. Hence it is her thus *submitting intentionally* to such nonsense – i.e., to not understanding – that manifests precisely the appropriate "*good will*" on our student's part in respect to our efforts to teach her.

The "picture" proposed by Weinberg's "definition" aims at building for our student and making her come to possess a kind of picture, or, rather, instill in her and make her imbued with some kind of intuition. Once the "definition" has gotten through, our student will be able to go on building intuitions *internal* to the conceptual (*and* experimental) workings of STR, intuitions of the sort physicists are basing themselves on and talking about all the time,[10] "intuitions [which don't arise *ex nihilo* but] are supplied by language itself" (TLP 6.233).[11] The "taut membrane without the membrane" will remain then the fleeting picture/no picture that bridged her way into STR and made her understand it. Hence the "ladder" in question, not only can, but should

[9] We will see in the third part of the paper how this captures the work that "showing" does in the *Tractatus*.

[10] Phrases like "I have the intuition that we should try to prove this rather than that" or "It is my intuition that the experiment will turn out this way rather than that" are common stock within the practice of physics. It lies outside the scope of the present paper to try explicating how such intuitions may resemble to or may differ from Kantian intuitions. However it is worthwhile to note that much of the work of Gaston Bachelard in *The New Scientific Spirit* (1934, 1975) and elsewhere aims at building the kind of intuitions we are talking about in respect to twentieth-century natural science and mathematics.

[11] In fact, Wittgenstein here refers such intuitions to mathematics. But we believe that the scope of his remark can cover situations like the one we are examining here while, in any case, physics is inextricably involved with mathematics since Galileo.

be “thrown away” (TLP 6.54) after the student has entered into the new paradigm of STR and understood the novel theory. She can and should throw it away when she is no more in need of visualizing or intuiting the electromagnetic field in inappropriate terms, i.e., in terms *external* to the physical theory in question.

To use a more offensive metaphor, the “ladder” we are talking about amounts to a kind of siege engine designed to exert pressure on the background “hinges” that leave no grammatical room to our student for understanding STR. The aim is to make her proceed to the appropriate change and replacement and the means is our judicious employment of nonsense. Pedagogical success in such a context amounts to our penetrating the student’s conceptual system, opening it up and taking it from the inside by flooding it with ours. Such success then makes the point we want to make emerge and stand sharply out of the background, inducing simultaneously to our student a “Eureka!” experience. This is exhibited by an exhilarating cry of “Yes, now I *see*!” by which our student means that she sees both the part of the world involved and the concepts she had been entertaining under a *new light*, the light provided by our own conceptual system, i.e., the conceptual system of STR. It is then, that is, that we have arrived to flood fully her system with ours. The old “hinges” have becomes dissolved and replaced by new, we have arrived to *show* to our student what she now *sees*, and thence her understanding can move freely within the grammatical space that has thus been opened up. Such understanding amounts to her “seeing the world rightly” (TLP 6.54) by the lights of STR.

We should add that our student’s coming to feel fully the effect of Weinberg’s “definition” on her is *sufficient* for her becoming clear *on all aspects* of the issue at hand.[12] This is to say that once the “picture” of the “taut membrane without the membrane” gets through, she sees in a flash and all at once (this is what the “Eureka!” experience amounts to) the concept we were asking her to understand, the “obstacles”[13] that had been preventing such understanding before – namely that (and how) the bedrock “adjustments” blindly allowing her understanding of CM were in fact leaving no room for her understanding the novel concept – and, thereby, why we were compelled to inflict nonsense upon her in order to make her open up the grammatical space capable, precisely, of hosting that novel concept. In other words, while *she* had *not* been talking nonsense at all when resisting us, for she was simply holding fast to the far from nonsensical CM, it was sufficient to feel the full force of *our* addressing to her the nonsensical “definition” in question in order to understand that this was *the only way* for us to get through to her and make her understand the contended concept.

We should stress once again that Weinberg’s elucidatory “definition” is not a bona fide proposition. Its nonsensical character precludes not only the possibility of its being deemed true or false but also the possibility of its being supported by arguments, explained or analyzed further in its own right or even explicated otherwise than through mere paraphrase. In the given teaching context, it either does the job it was

12 Of course, coming to feel *fully* such an effect presupposes a usually long struggle to come to terms with a whole host of associated issues; merely hearing or reading Weinberg’s definition is not enough.

13 We are referring to what Bachelard throughout his work calls “obstacles epistemologiques” These, we maintain, can be explicated in terms of the “silent adjustments” we have been talking about. We cannot undertake this work here.

designed for or it doesn't. This is to say that its utterance can only be considered as a kind of *action* performed within the teaching activity, as *a move of the strategy* we deploy to make our student understand the novel concept of the electromagnetic field. Its value can be assessed in this respect and in this respect only. This is to say that the only value it can bear is *effectiveness*: to the extent that it fulfils the purpose for which it was designed, it has worked as what we can call *telling* nonsense.

Of course, not all nonsense can be telling in this sense while there cannot be any hard and fast rules which could specify which particular piece of nonsense would be telling in any given, always singular, context.[14] Imagination, therefore, should necessarily be called upon in the ways we have tried to describe above. The need for such an appeal to the singularity of imagination is strengthened by the fact that not two students understand CM in exactly the same manner while blind spots in the understanding of both students and teachers usually exist, located at different places for different people. On the other hand, we should also stress that nonsense is rarely, if at all, totally pointless or altogether devoid of purpose. From the nonsense of Lewis Carroll to that of Baron Munchaußen, from the theater of the absurd to the "philosophical" novels of Borges, from the nonsensical languages invented by children during their games to the fundamental stipulation of psychoanalytic treatment, according to which one has to say whatever comes into mind, no matter how inconsequential, silly or absurd, nonsense has *a point* to make, *does* something to us, *affects* us in various ways. From this point of view, our dealings in colloquial language are largely dealings with nonsense. If we were not straitjacketed into the reigning "conception" of rationality, we could perhaps even risk the oxymoron: "in many real circumstances, our dealings with nonsense are full of meaning". The way nonsense affects us, however, becoming thus performatively meaningful is always heavily dependent on our actual relations to various linguistic, and not merely linguistic, contexts. No general category "nonsense" exists that can be divided in the abstract, *independently of context and purpose*, between telling nonsense and pointless or idle nonsense.

5. PART II: THE *TRACTATUS* ON PARADIGM CHANGE

In our efforts to understand how we can arrive to teach a radically novel scientific theory to somebody still entrenched in the old theory we helped ourselves to various remarks of the *Tractatus*. However such help might well prove coincidental and fortuitous. To see whether this is or is not the case or, which is the same, to attest whether the *Tractatus* can indeed be brought to bear on the issue at hand, we have to look closer at that work and ask whether and, if yes, in what particular way it countenances paradigm change in science.

To see how the *Tractatus* countenances paradigm change in science we start by remarking two things. First, that the *Tractatus* is concerned much more with logic than with grammar and, second, that, for its author at least, even the most radical grammatical changes, those that have successfully defied what appeared as self-evident beyond

[14] For an analysis of "nonsense" consonant with the present approach see Diamond (1995a).

all possible doubt (in the case at hand, that a wave which is *defined* as the propagation of a medium's disturbances may need no medium to propagate) cannot touch logic as such. Self-evidence is not the mark of logic or of logical propositions in general (TLP 5.4731, TLP 6.1271) and thus no act of imagination challenging self-evidence is in a position to impugn logic at all: logic "takes care of itself" (TLP 4.73) and thus always finds a way of reinstating itself. Moreover no act of imagination whatsoever can ignore or disparage logic; "logic fills the world" (TLP 5.61) and therefore even the wildest imaginative fancy cannot but remain, in this sense, worldly and hence logical: "It is clear that however different from the real world an imagined world might be, it must have something in common – a form – with the real world" (TLP 2.022).

It follows that the novel grammatical possibilities born by the kinds of wider grammatical space we have been discussing cannot be considered as implying novel *logical* possibilities. As "there can *never* be surprises in logic" (TLP 6.125, Wittgenstein underlines), no *novel* logical possibilities can rest dormant, waiting to be discovered by us; there can be no discoveries in logic. And in any case, logical possibilities, whatever they might be, cannot be abstracted away and be taken as lying "out there", passively in the offing. Regarding logical possibilities in this way would be placing logic outside language, thought and the world and putting it out of action. The conclusion is that to the extent that the *Tractatus* takes up grammar at all, the grammatical "level" cannot but remain distinct from the logical "level".

Thus the question becomes: does the *Tractatus* tackle grammar and all the associated issues we have been discussing? The answer, we take it, is positive despite what might appear on a first reading of that work.

To start with, there exists in the *Tractatus* the exact analogue of the grammatical "hinges" we have been talking about. These are "the enormously complicated silent adjustments[15] [allowing us] to understand colloquial language" (TLP 4.002). These "enormously complicated silent adjustments" intend to capture, we take it, precisely the workings of what Wittgenstein himself would later discuss as the grammatical "riverbed" of language (Wittgenstein, 1969, Sect. 97 ff.). Thus, if the "hinges" of that work intend to capture what remains *grammatically fixed* in a given context of linguistic usage,[16] then the "adjustments" of the *Tractatus* intend, symmetrically, to capture the *grammatical leeway* that permits our moving grammatically from one linguistic context to another. Thus in both cases Wittgenstein seems to be after what is commonly shared in the background by everybody partaking of any given linguistic community, even if his "hinges" tend to focus the examination at a given linguistic (and not merely linguistic) context while his "adjustments" tend to encompass the grammatical bedrock of colloquial language throughout.

[15] Ogden translates *stillschweigenden Abmachungen* precisely as "silent adjustments" while Pears and McGuinness translate it as "tacit conventions". Both renderings seem compatible with Wittgenstein's German, but we believe that the second translation can be profoundly misleading: it confers a conventionalist flavor on Wittgenstein's overall conception of language which, we believe, is nowhere supported in the *Tractatus*. In addition, "adjustments" marry very well with the "hinges" of *On Certainty*, Wittgenstein's latest writing. In addition, we believe that "silent" should be preferred to "tacit", for it resonates with Wittgenstein's 'injunction' in TLP 7: "… *muss man schweigen*".

[16] The late Wittgenstein would restrict such talk to a "language game", but this need not concern us here for such restriction does not seem to be at work in the *Tractatus*.

To restrict our terminology to that of the *Tractatus*, we can say, therefore, that to understand propositions in general, we are indebted "silently", i.e., without being aware of their existence and of their function, precisely to those "adjustments": it is these that give us the necessary leeway for understanding propositions of our linguistic peers which we may not have encountered as such ever before and which, moreover, are advanced in various and continuously changing circumstances or contexts. It is this leeway which permits us recognizing what is grammatical and what is not in any given circumstance; it is the same leeway which determines our distinguishing sense from nonsense and meaningfulness from meaninglessness in all the various contexts of actual communication. Our saying above that, in the passage from some paradigm to its successor, some grammatical "hinges" become removed, is therefore to say that some "*re*adjustments" have taken place in the "part"[17] of the "riverbed" which determines silently from the background the sense and the meaning of the concepts involved.

Given this, Wittgenstein's holding that "man possesses the capacity of constructing languages in which any sense can be expressed" (TLP 4.002) can be taken as pointing, among other things, at the very phenomenon of paradigm change, for "every sense" cannot but include "non classical" senses as well. This is no mere speculation on our part. That the intellectual upheavals brought about by the developments in physics – at least after the publication of Einstein's relativity paper in 1905 – could not but have touched him is manifested in TLP 6.341 and TLP 6.342. In these uncharacteristically discursive and lengthy remarks, Wittgenstein discusses, precisely, mechanics in terms of (conceptual) "networks" of different shapes ("triangular", "rectangular", etc.), expressly picturing or instantiating different possible theories of mechanics, like, we take it, CM and STR, that could "cover" the "surface" representing "the world", and of the relation those "networks" bear to logic. The remarks are nested within the family of propositions TLP 6.3–TLP 6.4 where Wittgenstein lays out his understanding of natural science.

These remarks highlight that conceptual distinctions, conceptual upheavals and the workings of the grammatical level that we have been describing are not Wittgenstein's prime concern in the *Tractatus*. Such matters engage only the "form of the clothes" (TLP 4.002) constituting colloquial language and have to do only with ways and means for arriving at a "finer", "simpler" or "more accurate" (TLP 6.342) description of the world. But exploring and elaborating such issues is the business of science and not that of Wittgenstein's logico-*philosophical* treatise. What is the affair of the *Tractatus* as a *logico*-philosophical treatise is that "the world … *can* be described in the particular way in which as a matter of fact it is described" (TLP 6.342, we are underlining) and no more. This is to say that it is not the object of that treatise what we may come to discover or to invent as a matter of fact, for "we must have to deal with what makes [such discovery or invention] possible" in the first place (TLP 5.555).[18]

17 The "riverbed" in question does not have proper parts; hence the quotes.

18 In fact Wittgenstein refers here only to forms in logic which could supposedly be invented. We believe, however, that his point can apply unproblematically to scientific discovery or invention in general.

Accordingly, to bring matters down to the level of purely *logical*, as opposed to grammatical, possibility the task Wittgenstein sets himself is to attain, at least in principle, "the one and only complete analysis" (TLP 3.25). At this level, each "elementary proposition" (TLP 4.221) consists simply of a "concatenation of names" (TLP 4.22) each of which stands for a "simple object" (TLP 2.02). This is to say that, at this level, the only level that is a matter of purely logical possibility, all the "enormously complicated silent adjustments" we have been discussing *have already been taken care of*. What happens at the level of concepts as well as the ways in which scientists handle the corresponding conceptual issues is of course, a different matter altogether, a matter concerning, as we have said, the "form of the clothes" constituting colloquial language.

It is in this sense that Wittgenstein lays to rest the whole set of issues involved in radical paradigm change in science. Conceptual upheavals like the coming to being of STR are the affair of science alone because science forms an autonomous enterprise following its own ways of proceeding. This is to say that science finds by itself and on its own how to overcome all kinds of obstacles that appear in its path as well as how to go on from there. Philosophy has not much to say on any of this for the very simple reason that the only "object" it can possibly have "is the logical clarification of thought" (TLP 4.112). "Philosophy is not one of the natural sciences" (TLP 4.111), it is "not a theory but an activity" (TLP 4.112), an activity which "consists essentially of elucidations" (ibid.). Accordingly, philosophy cannot and should not intervene in science and has no advice whatsoever to offer to scientists. Regarding the issue at hand, it has performed its elucidatory task in full once it has distinguished the level of grammar (that of the "silent adjustments" determining the forms of the "clothes" making up colloquial language) from the level of logic and, concomitantly, once it has shown that logic remains what it has always been even through the most radical paradigm shift. Once this point goes through, all philosophical worries attending paradigm change and, more generally, the workings of science should disappear.

An example of such a worry is the status of the "theoretical entities", i.e., the core of the debate between scientific realism and antirealism. From what we have said above it follows, we take it, that if indeed philosophy can have nothing substantial to say on science and its results then it is in no position to urge *scientists* to go in this or that way in that debate; the whole of it, as precisely a *philosophical* debate, would have been perfectly idle for Wittgenstein. The reasons for such an attitude should be obvious. Given science's autonomy, what the world is made up from (particles and fields, quarks and quasars or whatever) can be left safely to science's procedures even if – or rather because – science might very well change radically its mind on such issues. More strongly put, the fact that, on various occasions, science *did* change its mind in this respect (the existence of the ether seemed unshakable before the advent of STR or, conversely, quarks were explicitly introduced as useful fictions while today practically all physicists believe in their actual existence) displays clearly its autonomy, adds to its overall credibility and should put to rest all philosophical temptation to pontificate thereon. The totally unexpected advent of STR in Wittgenstein's time has shown once and for all that the world can in fact be described with "different [conceptual] networks" (TLP 6.341), that some such "networks" can be "simpler", "more accurate" and/or "*finer*" (ibid.) than others, that science possesses all the means necessary for adjudicating in the longer or shorter run such qualifications and that is that. There is nothing more to be said by philosophy on the relation between science and the world except to simply note along the way that "something is [indeed being] asserted about the world" (TLP 6.342) by the fact that it "can be described in the particular way in which as a matter of fact it is described" (ibid.). As we will attest in Part III, Wittgenstein's position is here even more dismissive: pretending to adjudicate issues like the existence of theoretical entities from outside science is to entitle oneself to a vantage point overarching all possible descriptions of the world that science can provide and, by the same token, a vantage point overarching the world as a whole. As we will explicate below, the goal of the *Tractatus*, at least on our reading, is to annihilate beyond possible appeal (i.e., logically) any entitlement whatsoever to such a vantage point.

The same kind of dismissal applies to the issue of relativism for reasons that are even clearer. We recall from Part I that the widening of the grammatical space attending a radical paradigm change is irreversible. The route leading back to conceiving the old concepts strictly in the way they were being conceived before the paradigm shift is blocked: the "assumption" disclosed cannot be forced to re-enter into the background. This is to say that the old paradigm has become *definitively* superseded and all scientific reasoning can take place only on the basis of the corresponding wider grammatical space. It follows that scientists are always residing within one or another paradigm and hence that no room is left at all for an extra-paradigmatic vantage point, i.e., a point from where one could impartially assess the relative merits and demerits of paradigms biased by none.

It is here that relativism shows its face: if we always find ourselves within a paradigm without the possibility of acceding to extra-paradigmatic neutral ground, it seems to follow that a paradigm is as good as any other and hence that we are free to choose the one that suits best our conventions, our interests, or even our whims. But on the above this is simply a *non sequitur*: the grammatical space available to the new paradigm is *objectively* wider than that available to the old or, appealing to the terminology of the *Tractatus*, the novel "network" is objectively "simpler", "finer" and/or "more accurate"[19] than the old. That the new paradigm has superseded *definitively* the old means that, after having undergone the relevant "Eureka!" experience, we *necessarily* reside within the novel grammatical space with no possibility of going back and hence no real choice between the two paradigms as such can be at issue. Theory comparison, theory choice, incommensurability and so forth can become matters of concern only *after* the novel paradigm has been established and *only* from the vantage point determined by the wider grammatical space while such concerns are, once again, the affair of science alone.[20] In this precise sense, objectivity need not imply neutrality in respect to paradigms and the attendant impartiality; their tie may appear unbreakable only

[19] As discussed in Baltas (2004), the idea of scientific progress can be "saved" on the basis of such objectively greater width. On the other hand, we should note that, after Kuhn, characterizations such as "simpler", "finer" or "more accurate" cannot stand as they are without qualifications. Incommensurability and all its associated issues should come into play. We repeat, however, that issues like these concern the conceptual level and do not form prime concerns for the *Tractatus*. On the other hand, the fact that Wittgenstein allows for (conceptual) "networks" of different *shapes*, as we have noted, might be interpreted as pointing at the incommensurability at play between different paradigms.

[20] For more details see Baltas (2004).

from a vantage point which overarches all paradigms and hence, since scientific paradigms talk about the world, a vantage point which overarches the world as a whole. Annihilating all philosophical temptation to occupy such an external vantage point – the main goal of the *Tractatus* on our reading – is at one and the same time annihilating the philosophical worry that radical paradigm change may imply relativism.

6. PART III: THE *TRACTATUS* AS A PARADIGM SHIFT: PREREQUISITES FOR A READING

In the first part of the paper we helped ourselves to some views of Wittgenstein's (both early and late) in order to understand the grammatical aspects of paradigm change in science and to explore the ways we should go about in order to teach the novel paradigm to students still entrenched in the old. In the second part of the paper we tried to show that, despite appearances and in its own very particular way, the *Tractatus* does countenance paradigm change in science. In what follows we will try laying out the fundamental prerequisites for doing, so to speak, the converse. We will be taking Wittgenstein as intending to convey a conception of philosophy radically different from any that a typical philosopher would expect or would be ready to accept. This is to say that, in respect to the typical philosopher reading his work, we will be taking Wittgenstein as occupying a position analogous to that occupied by the teacher of STR in respect to her student. This will allow us to exploit what we have been saying in order to shed some light on what the *Tractatus* itself is all about. Once we are allowed the requisite latitude in the use of the term, what we have just said amounts to our considering the *Tractatus* as proposing a radical "paradigm shift" for the whole "disciplinary matrix" of philosophy. The far from negligible differences between paradigm change in science and the particular "paradigm change" in philosophy which we are going to examine will be noted and discussed as we go along. *Our* reader, however, should not expect too much: if for no other reasons, space does not allow us to present more than the barest of outlines of this way of approaching the *Tractatus*.

The starting point of this approach, for which we will not offer any prior justification allowing what follows to speak for itself, is that the whole point of the *Tractatus* boils down to this: we are always and inevitably inside the world, inside thought and inside language and hence there cannot exist in the world, there cannot be articulated in thought and there cannot be formulated in language the coordinates of any kind of external Archimedean vantage point, allowing us to see, conceptually grasp, and talk about the world, thought and language from the outside in the philosophically typical all encompassing terms. Such an outside cannot exist, cannot be thought of and cannot be talked about. Resorting to a half forgotten philosophical characterization, we are maintaining, that is, that the whole point of the *Tractatus* is *immanence* and the task Wittgenstein has set himself is to show that immanence is the only *logically possible* philosophical perspective which, in addition, as TLP 6.54 has it, *necessarily self-destructs*.

Wittgenstein, however, does not appear, prima facie at least, to endorse what we have just been taking him to hold. Most of the propositions we have been helping ourselves to in what precedes appear as a bona fide propositions that can be variously argued for as well as employed for this or that purpose (e.g., for clarifying paradigm change in science the way we tried to do). And this despite the fact that many of them are obviously propositions issued from a vantage point appearing as capable of encompassing the world (or thought or language) as a whole, and hence a vantage point external to the world (or thought or language), fact which makes these propositions nonsensical on the immanentist view we are assuming Wittgenstein to hold.

But this need not deter us. Keeping firmly in mind that, on TLP 6.54, Wittgenstein does hold that all his propositions *are* nonsensical as the immanentist perspective would have it, we add at this juncture that just as it was not sufficient to simply expound STR in its own terms to make our student understand it, it may well be equally insufficient to propound the general coordinates of the immanentist perspective to make the typical philosopher forego her views to the benefit of ours. If the *Tractatus* is indeed after a major "paradigm shift" in philosophy, then to convey what it intends, its propositions must hook onto those of the typical philosopher by making the necessary concessions to her initially *not* understanding what is to be conveyed. And for hooking onto the typical philosopher in this way, the *Tractatus* has to issue the kind of propositions that would be expected from a philosophical treatise. It is only after the typical philosopher has been drawn in this way into the *Tractatus* that she can hopefully come to understand why the propositions of that work are formulated the way they are, how they all hang together and what their author intends to convey by them. Only after all this work is done can the reader experience the full force of those propositions and finally come to join the author's point of view by undergoing the relevant "Eureka!" experience. General pronouncements on the merits of the immanentist perspective are usually ineffective and remain perfectly idle.

If this holds water, the situation here is, in a sense, the inverse of the one we encountered when discussing paradigm change in physics. What we mean is that Weinberg's student advances sense while Weinberg's task is to demonstrate to her that a new grammatical possibility has to be taken into account. To achieve this, he issues openly nonsense. In the case at hand, the typical philosopher advances nonsense while, by helping herself to the illusionary entitlement of the existence of an external vantage point, she takes for granted that she advances sense. Thus Wittgenstein's task is to demonstrate that the typical philosopher advances in fact nonsense by showing to her that the entitlement in question is illusionary for it is ruled out by logic itself. Thus if Weinberg's task is to show that a novel grammatical possibility has to be taken into account, Wittgenstein's task is to show that an apparent grammatical possibility taken for granted is logically ruled out. To achieve this he has to imitate the procedures of the typical philosopher and thus appear to her as talking sense by seemingly helping himself too to the same illusionary entitlement. It is only in this way that the *Tractatus* can hook its reader and discharge its task in respect to her.

If elucidation and nonsense go hand in hand the way we have been describing, then Wittgenstein can elucidate to the typical philosopher that she helps herself to an illusionary entitlement only by judiciously employing telling nonsense. It follows that the nonsense Wittgenstein is employing is not profound or substantial nonsense, supposedly pointing at some deep "truths" beyond words.[21] Wittgenstein's nonsense has only *a point to make* which is to say that each proposition of his has a particular *task to perform*. Performing successfully this task comes to his effectively *showing* to the typical philosopher that – and how – she has helped herself to an illusionary entitlement, an entitlement completely ruled out by logic. This is all what Wittgenstein's "showing" amounts to. In other words, this "showing" is always and only a displaying or a putting on view that intends to make the typical philosopher come to *see* each time not some inexpressible "truth" or other, but *only some particular form or aspect* of the logical impossibility of entitling oneself to a vantage point external to the world and to thought and to language.[22]

To go on, we should underline another crucial difference between the position which Wittgenstein occupies in respect to his reader and the position which Weinberg occupies in respect to his student: in contradistinction to the fact that Weinberg *is based on the already existing* STR while his student is holding fast to CM, neither of which is, of course, nonsensical, there is no theory either Wittgenstein is targeting or that he can stand on in order to proceed with his tasks. Since for him "philosophy is not a theory but an activity" (TLP 4.112), no "body of doctrine" (TLP 4.112) whatsoever is either aimed at or can secure a foothold for the "ladder" he has to build, if he is to lead the typical philosopher out of nonsense. Both ends of the "ladder" in question either the "good" end, leading out of nonsense or the "bad" end that Wittgenstein is targeting cannot be suspended on anything. They hang in the void.

Given this, the question becomes: what kind of stuff can the "ladder" in question consist of if Wittgenstein is not proposing any circumscribable "body of doctrine" whatsoever? The answer is simple: this stuff consists of *nothing*, the reason being that our author has espoused the immanentist perspective from the very start. Aware that this requires clarification, we start by repeating that what defines the perspective in question is the claim that no external vantage point can possibly exist from where *any theory* about the world, thought and language in general can be formulated. Now, this claim is weird, not only because it is as nonsensical as Weinberg's definition (it rules out by what it is saying what it presupposes in order to convey what it says) but also because it precludes *any kind of assessment* (of its truth, of its value or whatever) from its outside. In other words, it is not only nonsensical but also *self-justifying*, refusing, by its very formulation, to submit itself to objective philosophical assessment; it concedes *absolutely nothing* to philosophical theory, the only instance whereby such an assessment could be forthcoming. It follows that the immanentist perspective cannot be stated or argued for with bona fide propositions. To espouse it means simply to *work it out* with no possibility of appealing to any theory describing how this work goes or prescribing how it should go. Such a working out can be carried out only on its own, in total disregard of any philosophical theory that would legitimize or guide it. In short, the philosophical activity for the immanentist perspective is *merely* this working

21 As, for example, Hacker (1986), among many others, would have it. For a fully blown criticism of that view, fully consonant with the present approach, see Conant (2002).

22 Examining the ingenious ways in which the *Tractatus* does manage to show what it intends lie outside the scope of the present paper.

out and nothing else. Hence *espousing the immanentist perspective and strictly considering philosophy as an activity are identical*.

This clarifies the kind of work that Wittgenstein does in the *Tractatus*. Since there is no theory whatsoever he intends proposing, working out the immanentist perspective can only amount to working on the views opposing this perspective (views entitling themselves in one way or another to the existence of the external vantage point at issue) *from within these views themselves*. Such work can only amount to gnawing at those views and chewing them up so as to finally arrive to eradicate them altogether. In less metaphorical terms, this is to say that Wittgenstein has to advance propositions that fasten themselves upon the propositions of the typical philosopher by imitating their grammatical structure and thus by appearing as typical philosophical propositions in their own right. But Wittgenstein's propositions fasten themselves upon the opposing propositions in the way that a parasite would fasten itself upon its host: their author's aim is not to counter them in a candid philosophical dialogue but to suck them, so to speak, within his whole elucidatory movement, so as to arrive to annihilate them all, together with his own parasitical propositions, by the end of the *Tractatus*. To achieve this as effectively as possible, he has, of course, to select and to arrange the propositions of the typical philosopher he will set himself up against or, which is the same, he has to order his own propositions according to their "logical importance" (TLP footnote).

In other words, Wittgenstein's goal is to demonstrate to the typical philosopher reading the *Tractatus* that the propositions she carries along while reading that work amount to idle or pointless nonsense. The means for showing this are the propositions of the *Tractatus* itself as they are ordered in respect to their "logical importance". This is to say that, insofar as the "ladder" constituting the *Tractatus* is effectively being climbed up to its last rung, these propositions function (or should function) as *performatively effective nonsense*. Excepting their performative function when put to work, Wittgenstein's propositions are no better than the propositions of the typical philosopher; "in themselves" they belong to the realm of mere nonsense and they return to it after their work is done and Wittgenstein's goal is accomplished. At the end of the day, the propositions of the typical philosopher have become fully erased, the eraser being Wittgenstein's propositions themselves, eraser which becomes fully consumed in completing its work.[23] Philosophical silence is thus attained.

This allows us to elaborate the distinction between telling nonsense and idle nonsense that we initially mentioned in Part I, for this, as we have seen, is precisely the distinction between the nonsense that Wittgenstein is using in full conscience from the nonsense the typical philosopher is using unawares. If philosophy is indeed an activity, then philosophy should necessarily have a *purpose*, as all human activities do, while there should be a distinction between good philosophy, the activity whose performance fulfills its purpose, and bad philosophy, the activity whose performance does not. An activity that fulfils its purpose is an activity that *works*, an activity that

[23] We can say, if we like, that the work done in the *Tractatus* aimed at rubbing out the Cheshire cat from inside it, as it were, not even leaving its smile to linger on. This is to say that Wittgenstein's work was undertaken in order to arrive precisely to the nothing of philosophical silence, a nothing very similar to that which Althusser's philosophy was consciously led to by its author. See Baltas (1995).

discharges the tasks it assigns to itself. An activity that does not fulfill its purpose is an aimless activity, an activity that remains *idle*, irrespective of what its practitioners might believe on the matter. But for Wittgenstein, the purpose of philosophy is that of elucidating thought (TLP 4.112). Hence good philosophy can only be the activity whose performance arrives to elucidate thought (always through nonsense) while bad philosophy can be only that whose performance does not arrive there or, more strongly put, an activity that finally confuses thought. Elucidatory nonsense is then telling nonsense, *performatively effective nonsense*, while confusing nonsense is *idle nonsense*, nonsense uttered without genuine aim or purpose, although its users may not be seeing this and believing the contrary. It should be added that the two cannot be distinguished by merely hearing or reading the signs at play, for the same sign can work differently in different contexts (TLP 3.321 and TLP 3.323). All this is to say that although nonsense does not express *a* thought by definition, it can nevertheless be *an instrument of the activity of thinking*, serving the purposes of elucidation, an instrument one can use effectively or ineffectively.

That nonsense can be used as an instrument of thinking allows us to understand what Wittgenstein says in the Preface of the *Tractatus*, regarding the aim of his work. "The book will draw a limit to thinking [i.e., to an activity] or rather … to the expression of thoughts. … The limit can, therefore, only be drawn in language and what lies outside the limit will be simply nonsense". What he means by this is, we take it, the following: clear and distinct (proper) thoughts, on the one hand, and what colloquial language can express, on the other, are *not* coextensive. A thought is a picture of reality (TLP 3) and language analyzed down to the level of "the one and only complete analysis" (TLP 3.25) should make this manifest. But our colloquial language "disguises thought" (TLP 4.002) and thus includes, by the same token, the possibility of nonsense. By *using this possibility judiciously as an instrument*, one may arrive to trace from the inside the limits to *thinking* (again: the limits to an *activity*) while thought remains always a picture of reality. What the limits to *thinking itself* look like from the inside Wittgenstein makes clear in the last sections of the *Tractatus* where he talks about the "mystical". Space, however, does not allow us to enter into this.

To the extent that the results of philosophical activity (either elucidatory or confusing) constitute nonsense, Wittgenstein's engaging it made him inevitably bathe in nonsense. The reason he took the plunge in the first place is relatively well rehearsed in the literature. Philosophical puzzles which cannot have an answer (TLP 4.003) crop up all the time tending to hold us captive. To get rid of such puzzles altogether, to cure himself from the "disease" of philosophy to which he had succumbed, but also to become vaccinated from all its possible recurrences, Wittgenstein decided to try elucidating once and for all how and why our engaging the philosophical activity amounts to our submerging in nonsense. The aim was to finally display the whole of philosophy as the nonsense it is and thus arrive alive and well to the Promised Land of *philosophical silence* (TLP 7) where the world appears spontaneously as it really is and hence all thoughts are crystal clear, where, in one word, one "sees the world rightly" (TLP 6.54).[24] This therapy, the *only possible therapy* if the immanentist perspective is the

[24] Diamond's *Realistic Spirit* is, I believe, precisely after this kind of "realism".

only philosophical perspective that is logically possible, had of course to pass through Wittgenstein's covering philosophy in its entirety. The *Tractatus* is the tangible result of precisely this exorbitant toil. It was Wittgenstein's considering that his toil had been fully successful after all that made its outcome seem to him "unassailable and definitive" (TLP Preface): all philosophical activity had been destroyed or blown up once and for all,[25] with no possibility of recovery or recombination.

After the work is done, after the whole of philosophy has been totally erased, there is nothing but silence because *there is literally nothing to be said* by philosophy, within philosophy or on philosophy while, of course, at one and the same time, *absolutely everything remains exactly as it is*, for philosophical activity in general cannot touch anything of substance regarding language, thought or the world. The philosophical activity in its entirety is and has always been perfectly futile, wheels within wheels turning idly in the void. This is exactly what the "truth" of Wittgenstein' system boils down to (TLP, Preface), a truth *indeed* "unassailable and definitive" to the extent that Wittgenstein's laborious *activity* has been fully successful *indeed*: despite philosophy's pretensions to the contrary, despite its understandable resistance to such total annulment, there is literally nothing for philosophy to say, there never was, there never will be. Wittgenstein may not have proved this last statement summarizing the endpoint of his toil, for there is no place for proof here; but he considered nevertheless that he had indeed managed to *perform* successfully the activity that *displayed* the "truth" of this statement beyond logically possible doubt.

7. CONCLUSION

Although Wittgenstein's later philosophy has found its way in philosophy of language, in philosophy of mind and even in philosophy of science,[26] the *Tractatus*, in contrast, has been left more or less standing alone, the object of study only of Wittgenstein scholars. The last part of the present paper should be perceived therefore as a gesture intending to remedy this situation. But there is more. The fact that what framed the present outline for approaching the *Tractatus* is, fundamentally, Kuhn's work on paradigm change implies two things. First, that the ways of approaching Wittgenstein's early work can indeed be various and hence that the lessons to be drawn from that work might prove far-reaching and, second, that Kuhn's work can be of value in unexpected philosophical regions, lying far outside philosophy of science proper.

The deeper reasons for this double possibility have already been hinted at. To the extent that Kuhn's work resolutely disallows any extra-paradigmatic vantage point, it can be considered as a work that espouses the immanentist perspective. To that extent it allies itself naturally to the *Tractatus*, if indeed, as we have been maintaining, the

25 Obviously such a hyperbolic claim cannot be accepted at face value; strong reasons should be forwarded in its support. To examine and to assess, however, why Wittgenstein could consider the outcome of his toil as fully successful requires our entering much more deeply into the *Tractatus* itself, something beyond the aims of the present paper.

26 The relations between Kuhn's work and Wittgenstein's later philosophy have been acknowledged by Kuhn himself and discussed in the literature. See Kindi (forthcoming).

whole point of that work does boil down to demonstrating the logical impossibility of the existence of any external vantage point in general. Our trying to bring together the two works can thus be hopefully beneficial to both: on the one hand, Kuhn's work on paradigm change can help us understand the *Tractatus* while the *Tractatus* can provide general philosophical grounds for Kuhn's work in philosophy of science and thereby find a kind of "application" in that field. If we have managed to convince our reader that exploring this relation might prove in the long run a worthwhile undertaking then the toil of writing this paper has been well spent.

BIBLIOGRAPHY

Bachelard, G. (1934, 1975) *Le nouvel esprit scientifique*. Paris, France: PUF.

Baltas, A. (1995) Louis Althusser: the Dialectics of Erasure and the Materialism of Silence. *Strategies*, 9/10, the Strategies Collective of UCLA, 152–194.

Baltas, A. (2004) On the Grammatical Aspects of Radical Scientific Discovery. *Philosophia Scientia*, 8(1), 169–201.

Conant, J. (1992) The Search for Logically Alien Thought: Descartes, Kant, Frege and the *Tractatus*. *Philosophical Topics*, 20(1), 115–180.

Conant, J. (2002) The Method of the *Tractatus.* In H. R. Erich (ed.) *From Frege to Wittgenstein*. Oxford: Oxford University Press, pp. 374–462.

Crary, A. and Read, R. (ed.) (2000) *The New Wittgenstein*. London: Routledge.

Diamond, C. (1995a) What Nonsense Might Be. In *The Realistic Spirit*. Cambridge: MIT, pp. 95–114.

Diamond, C. (1995b) Throwing away the Ladder: How to Read the *Tractatus*. In *The Realistic Spirit*. Cambridge: MIT, pp. 179–204.

Hacker, P. M. S. (1986) *Insight and Illusion*. Oxford: Oxford University Press.

Holton, G. (1988) *Thematic Origins of Scientific Thought*. Cambridge, MA: Harvard University Press.

Kindi, V. (forthcoming) *Kuhn and Wittgenstein: Philosophical Investigations on the Structure of Scientific Revolutions*. London: Routledge.

Kuhn, T. S. (1962) *The Structure of Scientific Revolutions*. Chicago, IL: Chicago University Press.

Ostrow, M. B. (2002) *Wittgenstein's Tractatus: A Dialectical Interpretation*. Cambridge: Cambridge University Press.

Weinberg, S. (1977) The Search for Unity: Notes for a History of Quantum Field Theory. *Daedalus*, 107, 17–35.

Weiner, J. (1990) *Frege in Perspective*. Ithaca, NY: Cornell University Press.

Wittgenstein, L. (1969) *On Certainty*. In G.E.M. Anscombe and G.H. von Wright (ed.). New York: Harper & Row.

Wittgenstein, L. (1986) *Tractatus Logico-Philosophicus*, C. K. Ogden (trans.). London: Routledge and Kegan Paul.

COMMENTARY ON "NONSENSE AND PARADIGM CHANGE", BY ARISTIDES BALTAS

ERIC OBERHEIM

Baltas's discussion of the positive functional roles of nonsense raises a number of interesting issues concerning the nature of conceptual progress. They have direct pragmatic implications for how best to teach new ideas. He aims to specify and describe a positive functional role of nonsense in learning and teaching, and then to argue that nonsense is a necessary component in the process of conceptual advance. He compares the role of nonsense in conceptual advance in the natural sciences to the kind of therapy Wittgenstein strived for in his later philosophy. His points are centered around a comparison of the elucidatory roles of Weinberg's nonsensical sentence:

> A field is a taut membrane without the membrane

and the infamous penultimate paragraph of Wittgenstein's *Tractaus Logico-Philosophicus*:

> My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless, when he has climbed out through them, on them, over them. (He must, so to speak, throw away the ladder, after he has climbed up on it.) He must surmount these propositions; then he sees the world rightly. (Wittgenstein, TLP 6.54)

Baltas does this in three acts. In the first act, he argues "the judicious use of nonsense *is indispensable* for elucidating the fundamental concepts of a radically novel scientific paradigm to students still entrenched in the superseded one" (Baltas, p. 49, italics inserted). By which he means that using nonsense "is the *only way* available for making the students enter the novel paradigm and come to understand its workings" (ibid., italics inserted). In other words, Baltas argues that nonsense *is necessary* for elucidating something that could not be understood otherwise. He uses the transition from Classical Mechanics to the Special Theory of Relativity to illustrate his conception of the relations between elucidation, nonsense and scientific advance. He reasons as follows: Relativistic concepts do not obey the grammatical rules governing Classical Mechanics. Consequently, the best way to elucidate them to students and proponents of Classical Mechanics is to *show* that the concepts cannot be understood from the entrenched, existing perspective. This is best done by talking a special kind of nonsense – by making a statement that "*precludes* by what it is saying what it

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 71–83.

presupposes in order to convey what it says" (Baltas, p. 56). If those entrenched in the older ideas appreciate the elucidation, then they realize that they will have to learn new kinds of concepts in order to understand according to Relativity Theory.

In the second act, Baltas argues that the *Tractatus* countenances Kuhnian paradigm change in science. He emphasizes that while Wittgenstein's *Tractatus* is mostly concerned with logic (which purportedly does not change according to Wittgenstein), it also deals with grammar and the possibility of grammatical change. He then claims that Wittgenstein's notion of grammatical change resembles Kuhn's notion of revolutionary conceptual change.

In the third act, Baltas tries to apply the results from the first two to explicate the infamous penultimate paragraph of Wittgenstein's *Tractatus* as exemplifying Wittgenstein's revolutionary conception of philosophy as a from of therapy. He makes an "elucidatory analogy" (Baltas, p. 50) between the role that talking nonsense purportedly plays in revolutionary advance in the natural sciences, and the purported role of nonsense in promoting the revolution in philosophy that Wittgenstein purportedly aspired to with the *Tractatus*. Baltas presents the *Tractatus* as an attempt to promote "the most radical 'paradigm change' in the whole 'disciplinary matrix' of philosophy, the 'paradigm change' aiming no less than to silence philosophy in its entirety" (ibid.). He claims that by writing the *Tractatus*, "Wittgenstein decided to try elucidating once and for all how and why our engaging the philosophical activity amounts to our submerging in nonsense" (Baltas, p. 68). According to Baltas, the *Tractatus* correctly teaches us "despite philosophy's pretensions to the contrary … there is literally nothing for philosophy to say, there never was, there never will be" (Baltas, p. 69). Then why did Wittgenstein write a philosophical treatise containing arguments in the first place? Baltas claims that Wittgenstein's "aim was to finally display the whole of philosophy as the nonsense it is and thus arrive alive and well to the Promised Land of *philosophical silence* (TLP 7) where the world appears spontaneously as it really is and hence all thoughts are crystal clear, where, in one word, one '*sees the world rightly*' (TLP 6.54)" (Baltas, p. 68). Baltas characterizes the aim of the *Tractatus* as a "therapy, the *only possible therapy* if the immanentist perspective is the only philosophical perspective" (Baltas pp. 68–69).

What to make of all of this?

Baltas appears to belong to what is called "The New Wittgensteinians": a loose group of intellectuals including Cora Diamond and James Conant (from whom he explicitly draws support). The New Wittgensteinians reject more orthodox interpretations of Wittgenstein's philosophical development (such as by E. Anscombe and P. Hacker). According to the New Wittgensteinians, even in the *Tractatus*, Wittgenstein was *not* primarily interested in *arguing* against other philosophers and their ideas. He was engaged in the kind of therapy he tried to use in his later philosophy:

> Wittgenstein's primary aim in philosophy is – to use a word he himself employs in characterizing his later philosophical procedures – a *therapeutic* one. (Crary, 2000, p. 1)

On this conception, philosophy is not an activity that attempts to advance metaphysical theories. Rather, the function of philosophizing is to help us work ourselves out of the confusions we become entangled in when philosophizing. While this has been a common interpretation of Wittgenstein's "later philosophy" (i.e., the *Philosophical Investigations*), the *New* Wittgensteinians are new in that they project this interpretation back into Wittgenstein's "early philosophy" (i.e., the *Tractatus*). Thus, Baltas argues that *in the Tractatus*, Wittgenstein attempted to show that:

> [A]fter the whole of philosophy has been totally erased, there is nothing but silence because *there is literally nothing to be said* by philosophy, within philosophy or on philosophy while, of course, at one and the same time, *absolutely everything remains exactly as it is,* for philosophical activity in general cannot touch anything of substance regarding language, thought, or the world. The philosophical activity in its entirety is and has always been perfectly futile, wheels within wheels turning idly in the void. (Baltas, p. 69)

The rest of this short commentary has two parts. In part 1, I argue against Baltas's claim that in the natural sciences, the *only* way to teach someone entrenched in an older paradigm the concepts of a newer paradigm is through the judicious use of nonsense. It is possible to teach new incommensurable concepts without talking nonsense. To help make sense of Weinberg's "nonsense", I introduce the distinction between propositional content and performative significance, and discuss the function of metaphor. I also offer some general critical remarks about Baltas's interpretation of Kuhn's model of scientific advance, and about his "elucidatory analogy" between Kuhn and Wittgenstein. Then in part 2, I will very briefly probe Baltas's interpretation of the penultimate paragraph of the *Tractatus*, on the basis of the characterization of the relationship between the *Tractatus* and the *Investigations* given in part 1.

(1) How can the proponents of a new paradigm teach those entrenched in the older ideas their new way of thinking? Baltas's answer is that we must speak a special kind of nonsense: We must make a claim that "*precludes* by what it is saying what it *presupposes* in order to convey what it says" (Baltas, p. 8). Before showing that it is possible to teach the concepts of a new theory *without* talking such nonsense, let us examine and develop exactly *why* Baltas believes that talking this special kind of nonsense is necessary for initiating paradigm shifts, and examine the *evidence* he presents on behalf of his claim.

Baltas chooses the notion of an "electromagnetic field" as developed by Special Relativity to argue that talking nonsense is necessarily involved in converting the adherents of the older paradigm to the new conceptual perspective. According to Baltas, since the development and confirmation of Maxwell's electromagnetic theory, physicists had conceived of a "wave" as "the propagation of a medium's disturbances" that "necessarily required the existence of the corresponding material carrier" that was called "the ether" (Baltas, p. 51). As all efforts to experimentally verify "the ether" failed, physics was in a state of crisis, which ended with Einstein's 1905 introduction of Special Relativity. *Failure of expectations* provided by the paradigm had set the stage for conceptual advance. According to Special Relativity, postulating the existence of the ether is not necessary for explaining the propagation of electromagnetic waves. This is because Special Relativity introduces a new concept: "the electromagnetic *field*, an entity altogether different from particles and media" (ibid.). Baltas proceeds to explain that the "most prominent characteristic of the electromagnetic field is precisely the fact that it can move by itself in vacuum without the support of any medium whatsoever" (Baltas, p. 51). According to Baltas, this concept of "field" is "entirely inconceivable" for classical mechanics. This is "simply because the electromagnetic waves that manifest the field's existence continue to be waves and a wave is classically *defined*, as we have just said, as the propagation of a medium's disturbances" (Baltas, p. 56). In other words, "the concept 'wave' is *analytically* related to the concept 'medium' " (ibid.). Suggesting the existence of a wave that propagates in a vacuum is like maintaining that some bachelors may be married males while defining bachelor's precisely as an unmarried male." (ibid.). This appears to be Baltas way of pointing out that the concept of a "field" does not fit into Classical Mechanics because they allow for action at a distance, something forbidden by the Newtonian, corpuscular worldview.

According to Baltas, the impossibility of introducing the concept of a field into Classical Mechanics creates a tension between Classical Mechanics and Special Theory of Relativity. It is eventually resolved as the concepts of the new paradigm takes hold, and the original "logical contradiction" is retrospectively reinterpreted as failure to notice logical possibilities (Baltas, p. 51). Baltas repeatedly uses a metaphor: the "hinges" that fasten a conceptual system to its background assumptions (borrowed from Wittgenstein's "On Certainty", Baltas, p. 52). Talking nonsense allows the "hinges" to swing open by loosening the conceptual connections between Classical Mechanics and the background assumptions the paradigm provides in the form of expectations. He claims that these "hinges" are represented in the *Tractatus* as the "enormously complicated silent adjustments" that must be made in order to facilitate a paradigm change (see Wittgenstein, *Tractatus*, 4.002, cited in Baltas, p. 60). He uses these ideas to describe the idea that new background assumptions about "waves" were needed in order to make the new concept of "field" intelligible to those entrenched in the older views:

> Removing the old "hinges" is equivalent to widening the grammatical space available to the inquiry. This wider grammatical space can host a *re*interpretation of the old conceptual system as well as, in some cases like the one at hand, an imperfect rendition of parts of the old conceptual system in terms of the new. Hence, within the bounds set by the novel conceptual system, these parts – precisely as reinterpreted – can sometimes still remain useful to scientific practice [such as when "the relevant parts of the novel conceptual system possess a 'classical limit'", inserted from the footnote to this passage]. However, the reverse movement is blocked, for the old, narrower, grammatical space cannot accommodate what the new allows. This makes the two succeeding paradigms asymmetrical while the incommensurability between the two, in Kuhn's (1962) sense of the term, is based on this asymmetry. (Baltas, p. 53)

According to Baltas, incommensurability is based on an asymmetry in the possible conceptual spaces of Classical Mechanics and Special Relativity. Baltas emphasizes that the transition in conceptual space is permanent and irreversible. He suggests that it is impossible to return to the old narrower conceptual space:

> After the conceptual tumult of paradigm change has settled down and the new conceptual system has become established, the wider grammatical space subtending it has become instituted *for good*: on the one hand, all further inquiry is based on it while, on the other, returning to the old, narrower, grammatical space becomes impossible.… Such widening of the grammatical space is therefore an *irreversible* achievement and hence theory choice and all the attendant issues can be negotiated only on the basis of the novel grammatical space. (Baltas, p. 53)

According to Baltas, because the two paradigms are incommensurable, "proof, evidence and all that are *insufficient*, that is, *per se powerless*" (Baltas, p. 55) to make those entrenched in Classical Mechanics aware of the superiority of Relativity Theory; and consequently, "it would be impossible to make our student [of classical mechanics] understand us through our simply expounding STR and the evidence for it" (ibid.). As rationality and proof are "powerless" in this situation, we need to "build a kind of bridge … which would enable us to meet [her] half-way" (Baltas, pp. 55–56). This bridge necessarily involves talking nonsense, because:

> This is a bridge or "ladder" which, to be precisely a means of connecting us two, has to share elements from both systems, employ terms, locutions and pictures which are somehow common to both. Since each system is self-consistent while the two are asymmetrical… the bridge or "ladder" connecting the two *cannot but constitute nonsense* from the point of view of either. Elucidation and nonsense thus seem to go hand in hand. (Baltas, p. 56, italics in original)

So according to Baltas, we cannot simply use the new concepts to teach the new concepts. We need to build a conceptual bridge by talking nonsense that elucidates the new conceptual possibilities. Baltas illustrates his case for the necessity of talking nonsense with a single example. He cites a claim Weinberg made when trying to explain Special Relativity to students of classical mechanics: "A field is a taut membrane without the membrane" (Weinberg, 1977, cited in Baltas, p. 56). Baltas claims,

> [T]his obviously self-destroying 'definition' provides exactly the kind of bridge or "ladder" we are after: its first part makes reference to CM (taut membranes are legitimate objects of that theory) while the second (that the membrane in question, like the ether, does not exist) sits squarely within STR. (Baltas, p. 56)

The definition is self-contradictory, and therefore nonsense, and it exposes the incommensurable gap between the two theories:

> [T]he 'definition' *should* destroy itself and should destroy itself *thoroughly*, to the point of becoming fully nonsensical. In this way and only in this way its very failure becomes the condition for its success. (Baltas, p. 57)

According to Baltas, Weinberg's statement is not part of a rational argument. Rather, it *exhibits* the impossibility of understanding Special Relativity in terms of Classical Mechanics, and it is *only* by showing or exhibiting that it is possible to come to understand Special Relativity's concept of the electromagnetic. It is only by talking nonsense that we can "pave the way" (ibid.) for properly understanding the new concept. Thus, "elucidation and nonsense go hand in hand" (Baltas, p. 56). In sum, Baltas believes that talking nonsense is necessary for converting those entrenched in the concepts of an older paradigm to the newer point of view, because without talking nonsense, we cannot demonstrate the impossibility of learning the new theory from the older point of view. So you have to talk nonsense to show what cannot be said. The nonsense required is sentences that combine concepts from incommensurable frameworks, bridging the gap between them.

What exactly makes some sentence nonsense? Baltas identifies a specific kind of nonsense and the necessary roles that he thinks it plays in conceptual advance and teaching new ideas, but he offers no characterization of a meaningful statement, and he provides no criteria by which we can distinguish nonsense from statements that have sense. He tells us that no general category "nonsense" exists independently of context (Baltas, p. 59). In his *not* defining "nonsense", he does "risk" the "oxymoron: in many real circumstances, our dealings with nonsense are full of meaning" (ibid.). Baltas also claims that "nonsense does not express *a* thought by definition, it can nevertheless be *an instrument of the activity of thinking*, serving the purposes of elucidation, an instrument one can use effectively or ineffectively" (Baltas, p. 68, italics in original). So perhaps, nonsense is sentences that do not express a thought. But Weinberg's "definition" seems to express some thought, even if it may involve relating incompatible concepts in a way that breaks the grammatical rules (or "logical syntax") of the languages to which they belong. (By the way, the concepts of incommensurable theories do not "logically contradict" one another, as Baltas suggests. Incommensurable theories have no logical relations. The theories are mutually exclusive. Their concepts are logically incompatible, not logically inconsistent.) Baltas seems to suggest that nonsense does not express a coherent thought. But then, which thoughts are coherent and which are not? Baltas sees a lot of nonsense. He mentions that of Lewis Carroll, Baron Münchhausen, absurd theater, the philosophical novels of Borges, and the propositions in Wittgenstein's *Tractatus*, and Weinberg's purported self-contradictory "definition" of a field. For him, "nonsense" is a very big term.

In order to get a grip on Baltas's claim about the necessity of talking nonsense, and the positive function of talking nonsense in scientific advance and learning and teaching science, we need a more specific notion of sense, not just examples of nonsense. There is one near to hand. In the *Tractatus*, Wittgenstein develops a technical notion of sense and nonsense, according to which the propositions of a formal language have sense only if they describe a possible state of affairs that could have been otherwise. In the *Tractatus*, the truth of a proposition is determined by how things are. Logic is concerned with the prior question of which propositions are capable of representing reality in the first place. Consequently, in the *Tractatus*, the sentence "2 + 2 = 4" has no sense, because it could not be otherwise. Well-formed sentences of arithmetic are not propositions with meaningful content. They are part of the syntactical rules that determine which sentences are well-formed in the first place. They express the rules of a language. Strictly speaking, they are nonsense. If we use Wittgenstein's early conceptions of sense and nonsense, then it seems obvious that scientists have to use nonsense to develop and teach their theories all the time, and Baltas's claim would become trivial.

Is Weinberg's claim about fields really intended to be a *definition*? Baltas always puts the term definition, as applied to Weinberg's sentence, in scare quotes. He is aware of the fact that Weinberg was not actually offering a literal definition of a field. If it is not a definition, then what is it? Perhaps he was making a joke. It is difficult to say without the relevant context. Baltas has not provided the page numbers to Weinberg's purported nonsensical "definition". This makes it very difficult to assess its intended function by checking its context. But if there is one lesson in Wittgenstein's *Investigations* relevant here, it is that context is often necessary for understanding meaning.

Instead of calling Weinberg's sentences a nonsensical definition, we might treat it as a *speech act* (to draw on Austin's vocabulary, see Austin, 1962). On this view, to say something is to do something. For example, to make the statement, "I promise to be nice", is to perform an act of promising, not to make a true or false statement. If we distinguish propositional content from performative significance, then the mysteriousness about the positive function of nonsense disappears. For example, in the *Investigations*, it says that if a lion could talk we would not understand it (Wittgenstein, 1958, p. 223[e]). He may have wanted to use this aphorism to emphasize certain features of language, not to make a literal statement about lions. (On Wittgenstein's own later account of meaning as use, one might argue that lions do talk, and we do understand them.) The point of this aphorism may not be its propositional content. The point of the aphorism may be its performative significance: perhaps to try to emphasize the role of empathy in understanding, or to emphasize the need to share a form of life in order to communicate. Context is not necessary to ascertain the propositional content of a sentence. It is necessary to ascertain the performative significance. With this distinction, we can redescribe Baltas's point: He is emphasizing the performative significance of Weinberg's sentence, and rejecting its propositional content. This distinction does not only help elucidate the functional role of nonsensical sentences, it also bears directly on one of the main differences between the *Tractatus* and the *Investigations*. The *Tractatus* takes the world to be everything that is the case. A working assumption is that the main function of formal languages is to express true or false propositions. In this way, natural language should reduce to formal language. Performative significance is reduced to propositional content. By contrast, in Wittgenstein's later philosophy, a main function of natural language is not just to convey propositional content, but to get people to act in certain ways. Recall the example at the very beginning in Sect. 1 of sending someone shopping with a slip marked "five red apples" (Wittgenstein, 1958, p. 2[e]). Language has different functions in different contexts. I might say "The door is open" to imply that someone ought to close the door, not to state an obvious fact about the door. Using language involves more than conveying propositional content. It has performative significance.

So perhaps Weinberg's sentence is not supposed to be a special kind of nonsensical "definition" that "*precludes* by what it is saying what it *presupposes* in order to convey what it says" (Baltas, p. 56). Perhaps it is just a metaphor. If a field is a taut membrane without the membrane, then a field is just taut emptiness. Perhaps that was Weinberg's point. Is the sentence still nonsense? Baltas uses an awful lot of metaphorical language to make his point about the necessity of talking nonsense. For example, he talks of the "hinges of understanding", "penetrating" and "flooding" bodies of thought, "bridges" and "ladders" of nonsense, "bathing in nonsense", "arriving at the Promised Land of philosophical silence", etc. Is it all nonsense? If, like in the case of using arithmetic, using a metaphor is talking nonsense, then talking nonsense is certainly part of the natural sciences, and Baltas's claims would just be a very nonsensical way of making a rather trivial point: Scientists bend the rules of language to expand their conceptual resources. However, metaphors are not really nonsensical. They simply exploit the normal meanings of our words in order to make a pragmatic point. By saying something that is obviously wrong in its literal sense or obviously inapplicable to the situation, they force the hearer to reevaluate the statement in a non-literal sense in order to understand what the performative significance of the utterance might be

(See Chap. 2 of Grice (1989) and Essay 17 in Davidson (1984)). It seems to me that Weinberg's sentence is just such a metaphor. It is a metaphor that should be treated as a speech act with performative significance, not a nonsensical "definition" or pseudo-proposition that "precludes by what it is saying what it presupposes in order to convey what is says" (Baltas, p. 56).

Baltas calls this "definition" a "picture" that should be thrown away once it has served its purpose (Baltas, pp. 57–58). He stresses that Weinberg's "elucidatory "definition" is not a *bona fide* proposition" (Baltas, p. 58). By combining concepts from incommensurable conceptual frameworks, Weinberg's nonsense breaks the syntactic rules necessary for ascribing a truth-value to it. According to Baltas, its nonsensical character precludes not only the possibility of its being deemed true or false but also the possibility of its being supported by arguments, explained or analyzed further in its own right or even explicated otherwise than through mere paraphrase" (Baltas, p. 58). But Baltas has tried to explain and analyze Weinberg's sentence (perhaps not in its own right?). He even specified the kind of nonsense he is talking about: a sentence that *precludes* by what it is saying what it *presupposes* in order to convey what it say. But I should confess that this explanation is not very clear to me. How exactly does Weinberg's sentence *preclude* by what it is saying what it *presupposes* in order to convey what it say? If I say, "I am going to the train station to pick up my sister", then it makes sense to say that by doing so I presuppose things like that I am not an only child and that I have a sister. What exactly is Weinberg's sentence supposed to be presupposing? Does it presuppose that membranes can be taut? I am not exactly sure what Weinberg's sentence presupposes (if anything), nor how it might be precluded by what the sentence says, nor how that should contribute to what it is saying.

Moreover, there seem to be easier ways to explain the sentence. I already gave one example above. Treat it as a metaphor. But there are also other ways that I can try to explain it so that it is no longer nonsense. I might change Weinberg's sentence by adding just one word: "A field is *like* a taut membrane without the membrane." Now it is no longer a "nonsensical", "elucidatory", "picture", "definition". It is simply an analogy. Could we use this analogy to exhibit the impossibility of providing a strict definition of the notion of field, as it is used in Relativity Theory, using only the conceptual resources of Classical Mechanics? Why not? Making analogies is a fine means of introducing students to new ideas. Are analogies and metaphors better or worse than nonsense (in Baltas's sense) for exhibiting that new concepts that do not fit within the conceptual framework of Classical Mechanics are needed to understand according to Relativity Theory? Who knows. But to support his claim that talking nonsense is necessary for teaching the concepts of new paradigms, Baltas needs to rule out all other possibilities. He has not even considered any. What if instead of this "elucidatory definition", Weinberg had simply said, "It is impossible to define the relativistic notion of a field with classical concepts because the waves that manifest the field's existence move through a vacuum without needing a medium, which is analytically impossible given the classical definition of a wave as needing a medium through which to propagate." Is that nonsense? I do not think so. If not, why wouldn't that work to convey to students or scientists entrenched in Classical Mechanics the fact that they will need to redefine their terms to understand the new theory?

It may be true that defining the concept of a field as it is used in Relativity Theory is impossible within the classical framework, but do we really need a classical "definition" of a concept from relativity to build a nonsensical bridge between the two paradigms? Kuhn argued that empirical terms get meaning through learning how to apply them to concrete problems within the theoretical context to which they belong. For example, to understand the term "mass" in the Newtonian sense, one needs to understand how mass relates to acceleration, and how one can use "mass" to describe the motion of a pendulum. With the advent of Relativity Theory, we have new concepts of "mass". Do we *have* to learn this new language by trying, and failing, to translate from the older language, producing nonsense? Or can we simply learn the new language in its own terms, temporarily putting aside Classical Mechanics as an inappropriate tool for addressing certain problems. Baltas simply claims that this latter possibility is impossible. He gives no argument for this controversial view.

Baltas argues that because the two theories are incommensurable, it would be impossible to make students understand the new theory simply by expounding it and the evidence for it. He reasons that we cannot use "proof, evidence and all of that" because we do not share common background assumptions, and when proof is thus powerless, it seems that we can resort only to elucidations. Is there really such a strict dichotomy between logical proof based on evidence and elucidation through talking nonsense? Baltas seems to imply that either we can use proof, or we have to resort to talking nonsense – that there is no middle ground. *But it was exactly Kuhn's point that the rationality of the natural sciences is less rigorous than that of formal logic, without thereby becoming irrational.* Similarly, Feyerabend introduced the notion of incommensurability to expose the inadequacy of logical empiricist models of scientific rationality, not to argue that science was irrational.

Kuhn did argue that theory choice cannot be *unequivocally* settled by logic and experiment *alone* (Kuhn, 1970, p. 94), and he did emphasize the roles that well-known epistemic values play in comparing theories. These include simplicity, universality, fertility, and beauty. Empirical support is an important epistemic value, but empirical fit is a matter of degree, and it is but one epistemic value among many used to compare the relative merits of competing theories. However, Kuhn equally emphasized that he never meant to suggest that logic and experimental findings are irrelevant to theory choice (see Hoyningen-Huene, 1993, p. 244), or that theory choice is irrational because of incommensurability. After the invention of a radically new theory, communication may at first be only partial, as scientists do not always carefully explain which concepts they are using and how they differ from more orthodox views. But it need not be. On Kuhn's account, scientists can use arguments that *persuade* students of the older paradigm to adopt the new concepts, even though these arguments are never *absolutely* definitive, like a sound logical proof. With the notion of incommensurability, Kuhn's point was *not* the impossibility that the two theories can be rationally compared, but that there is the possibility for *rational disagreement* between the proponents of two paradigms. Similarly, Kuhn's point about the impossibility of translating incommensurable theories does not have the consequence that scientists have to talk nonsense, but simply that scientists do not have to translate their theories into each other in order to understand. They can *rationally* compare them in terms of degree of accuracy, scope, simplicity, fertility, universality, internal coherence, over all coherence with other theories, etc., without translating them into each other. That theory comparison involves making judgments about the relative weighting of epistemic values is not even a particularly Kuhnian idea. It had been discussed in detail by Duhem in 1906.

Baltas repeatedly treats Kuhnian conceptual revolutions in the natural sciences as an analogue to the idea of "widening" (pp. 3–6, 14, 15) or "enlarging" (p. 6) the grammatical space that he sees in the *Tractatus*. Kuhn specified his point about the nature of conceptual change by emphasizing taxonomic changes that occur in scientific revolutions. New theories group objects into different natural kinds. And Baltas is right that this is somehow reminiscent of Wittgenstein's discussion of "enormously complicated silent adjustments". Kuhn's account of taxonomic change can indeed be understood as a specification of a form of grammatical change in the Wittgensteinian sense. Changing how objects are collected into kind terms alters logical syntax. But is this really an appropriate analogue to the "widening" or "enlarging" of grammatical space in Wittgenstein's sense? These metaphors may be highly misleading as "elucidatory analogies" in some important respects. By calling two theories incommensurable, Kuhn was emphasizing the point that the new concepts are *exactly not* simply an extension of the older conceptual system. We cannot simply enlarge the grammatical space by adding the concept of field into the older conceptual framework – somehow reinterpreting the older ideas. Rather, the older framework is altogether dropped for the purpose of understanding according to the new framework. (It may still be used instrumentally, on in other contexts.) By contrast, Baltas suggests that the "wider grammatical space can host a *re*interpretation of the old conceptual system as well as, in some cases like the one at hand, an imperfect rendition of parts of the old conceptual system in terms of the new" (Baltas, p. 53). But by emphasizing the local holism of language together with the conceptual replacement involved in revolutionary advance, this is just the kind of account that Kuhn was rejecting. According to Kuhn and Feyerabend, the new conceptual framework does not "enlarge" by "widening" the older one. It *replaces* it. That is why progress through revolution is not cumulative. The new conceptual framework need not be "wider" in the sense that it explains everything the old system could explain, plus something extra, i.e., the anomaly to the older system – as the metaphors of "widening" or "enlarging" may seem to suggest. Quite to the contrary, both Kuhn and Feyerabend emphasized that the new conceptual space may be "narrower" in the sense that it may be unable to explain some of the relevant phenomena that had been successfully explained by the superceded theory. This is even called "Kuhn loss" (see, e.g., Bird, 2005). (Feyerabend also repeatedly emphasized the point that revolutionary progress often involves accepting some explanatory loss.) According to Kuhn, scientific revolutions do involve gains in problem solving capacity. But these gains come at the cost of losses in the ability to explain certain phenomena still deemed to be relevant, and at the cost of the loss of specific scientific problems. This can result in the narrowing of the field of research, which in turn can result in increased difficulty in communicating with outsiders due to increased specialization (see Hoyningen-Huene, 1989, pp. 260 ff.). For these reasons, the metaphors of "widening" or "enlarging'" grammatical space seem inappropriate to Kuhn's account without some qualifications.

Baltas also suggests that the required bridge of nonsense is one-way, implying that once someone has learned the concepts of a new paradigm, there is no going back: "That the new paradigm has superseded *definitively* the old means that, after having undergone the relevant "*Eureka!*" experience, we *necessarily* reside within the novel grammatical space with no possibility of going back and hence no real choice between the two paradigms as such can be at issue" (Baltas p. 63). But Baltas has Kuhn wrong. Kuhn suggested that in fact scientists rarely do go back, because of their interests, but how would Kuhn's job as a hermeneutic historian of science even have been possible if it were impossible to go back to understanding older theories in their own terms. That is just how Kuhn claimed to have discovered incommensurability in the first place (see Hoyningen-Huene, 1993, p. 41; Oberheim, 2005, pp. 365 ff.). Are we really to believe that having developed Relativity Theory, Einstein was no longer capable of understanding the concepts of Classical Mechanics in their own right? Or that having understood Statistical Thermodynamics, we can no longer use Classical Thermodynamics to make calculations about large systems? Or that it is impossible for us to imagine the geocentric worldview just because we usually use a different concept of planet that does not include the sun? The point of incommensurability is that one cannot use the concepts of two incommensurable systems *at the same time* – not that one can never go back.

Using the concepts of incommensurable frameworks in a single sentence does create nonsense. Such sentence are not well-formulated according to the grammars (in Wittgenstein's sense) of either of the theories to which the concepts belong. Moreover, Baltas's makes a very good point that such nonsensical sentences can still be used to make a point. But Baltas has not shown that scientists and teachers *must* do this. Baltas should explain why scientists cannot learn new theories in their own language from scratch (without building bridges of nonsense). He should qualify his "elucidatory analogy" between Kuhn and Wittgenstein so that it is less misleading, and he should explain why scientists cannot move back and forth between different incommensurable frameworks as easily as they can change their shirts.

(2) Baltas, like the other New Wittgensteinians, see the aim of the *Tractatus* as elucidating the complete futility of all philosophy, and he basis his interpretation on the penultimate paragraph. But as is the case with Weinberg's sentence, here too there is an easier way to understand the penultimate paragraph. Use the relevant context!

On one reading, the reason Wittgenstein called the propositions of the *Tractatus* senseless is *not* that the *Tractatus* establishes the futility of all philosophical reasoning. In fact, on this reading one substantial aim of the *Tractatus* was to challenge Russell and Frege's interpretations of logic. Frege thought that logic is a science. It discovers fundamental truths about the relationships between propositions that are universal. They are true of all thoughts. On his conception, the results of a logical investigation would be the general laws of thinking. Russell, on the other hand, thought that logic is concerned with the description of the most general features of the universe (For a discussion, see, e.g., Hacker, 1996, pp. 26–29). Wittgenstein thought that both of these interpretations of logic were incorrect, and he offered a new one. Logic is not descriptive in the way empirical claims are. Empirical claims are contingent. They state states of affairs that could have been otherwise. Logic states the syntactical rules that are necessary to determine whether a sentence is a well-formed proposition in the first place. They do not have a truth-value like empirical claims. Perhaps, the reason that Wittgenstein calls the “propositions” of his treatise senseless is that according to the technical definition of sense that he develops as part of the new interpretation of logic delineated in the *Tractatus*, the propositions contained in the *Tractatus* are, strictly speaking, senseless. They are not pictures of possible states of affairs that could be otherwise. With the *Tractatus*, Wittgenstein intended to *show* the limits of meaningful discourse. Wittgenstein reasoned that we use thought to represent reality and thoughts can be completely expressed in language. He viewed logic as a necessary precondition for thought, and he interpreted logic as setting limits on thinking by setting limits on the linguistic expression of thought. The system of rules that determine whether or not a string of expressions is a meaningful proposition (“logical syntax”) cannot themselves be expressed in meaningful propositions, because they would state *necessary* properties of language; and consequently, they would not express possible states of affairs that could be otherwise. For this reason, they cannot be *said*, but only *shown*. (They have performative significance, not propositional content.) Once we have been shown the point, and learned the lesson about meaning and logic, we can throw it away like a ladder we no longer need. That is the reason Wittgenstein called his propositions senseless toward the very end of the *Tractatus*. And once we understand this point, we no longer need the *Tractatus* to make it. It is not a very mysterious point, and we do not need a strained analogy to some insights about the purportedly necessary use of nonsense in paradigm change in the natural sciences to understand it.

Perhaps this interpretation of the penultimate paragraph of the *Tractatus* is incorrect. But before we turn to a rather stretched elucidatory analogy concerning the positive functional roles of nonsensical sentences in scientific advance for help, and before we begin projecting Wittgenstein’s later views about the aim of philosophy into his earlier texts where they do not belong, we ought to try to exhaust the resources within the *Tractatus*, especially something as relevant as Wittgenstein’s explication of sense and nonsense in the *Tractatus*, which Baltas has not even mentioned. Baltas owes us an explanation of why this simpler explanation of the penultimate paragraph of the *Tractatus* fails.

In concluding, Baltas suggests that “despite philosophy’s pretensions to the contrary, despite its understandable resistance to such total annulment, there is literally nothing for philosophy to say, there never was, there never will be” (Baltas, p. 69). If Baltas really believes that there is, and never will be, anything for philosophy to say, that all of philosophy is nonsense, then why did Baltas write an article trying to use some of Kuhn’s ideas to explicate or elucidate some of Wittgenstein’s. The answer seems obvious. Baltas appears to be attempting to use nonsense to elucidate the futility of propounding theses in philosophy in order to encourage a New Wittgensteinian revolution. In so doing, Baltas has indeed succeeded in making me feel as if I had been “bathed in nonsense”, but he has not brought me to “The Promised Land of *philosophical silence*”, where all my thoughts are crystal clear, where the world “spontaneously appears as it really is” – free of philosophical confusion. Baltas should not be discontent with my characterization of his text as “nonsense”. After all, on his own account, nonsense is supposed to be the first necessary step to accomplishing the New Wittgensteinian revolution in philosophy he supports. Perhaps a more appropriate commentary would just have been silence.

Acknowledgements

I would like to thank Julia Staffel for helpful suggestions on a draft of this comment – especially, but not limited to, some insightful discussions on the roles and functions of metaphor.

BIBLIOGRAPHY

Austin, J. L. (1962) *How to Do Things with Words*. Oxford: Oxford University Press.

Baltas, A. (2008) Nonsense and Paradigm Change. This volume, pp. 47–68.

Bird, A. (2005) Thomas Kuhn. In E. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Online.

Crary, A. (2000) Introduction. In A. Carery and R. Rupert (eds.) *The New Wittgenstein*. London: Routledge, pp. 1–18.

Davidson, D. (1984) *Inquires into Truth and Interpretation*. Oxford: Clarendon.

Grice, P. (1989) *Studies in the Way of Words*. London: Harvard University Press.

Hacker, P. (1996) *Wittgenstein's Place in Twentieth-Century Analytic Philosophy.* Oxford: Blackwell.

Hoyningen-Huene, P. (1993) *Reconstructing Scientific Revolutions. Thomas S. Kuhn's Philosophy of Science*. Chicago, IL: Chicago University Press.

Kuhn, T. (1970) *The Structure of Scientific Revolutions*, 2nd ed. Chicago, IL: Chicago University Press.

Oberheim, E. (2005) On the historical origins of the contemporary notion of incommensurability: Paul Feyerabend's assault on conceptual conservativism. *Studies in the History and Philosophy of Science*, 36, 363–390.

Wittgenstein, L. (1958) *Philosophical Investigations*, E. Anscombe (trans.). New York: Macmillan.

Wittgenstein, L. (1969) *On Certainty*. In G. E. M. Anscombe and G. H. von Wright (ed.). New York: Harper & Row.

Wittgenstein, L. (1986) *Tractatus Logico-Philosophicus*, C. Ogden (trans.). London: Routledge and Kegan Paul.

PART 3

INTRA-THEORETICAL CHANGE, AS A SUBJECTIVE CREATIVE ELUCIDATION OF AN OBJECTIVE FORMERLY PRESENT CONTENT

FROM ONE VERSION TO THE OTHER: INTRA-THEORETICAL CHANGE

ANOUK BARBEROUSSE

Abstract Describing scientific change involves displaying scientific theories and their content. This is done by way of (rational) reconstructions, which usually neglect the variety of versions in which a scientific theory may appear. Although they possess a unity allowing them to bear the same name for decades, or even centuries, theories are diachronic, evolving entities. How should their evolution, i.e., intra-theoretical change, be described? I claim that this description is best done in terms of a series of conceptual undertakings on the part of scientists aiming at examining the meaning of theoretical principles and concepts and broadening their domain. I present the implications of this hypothesis, which follows Robert Brandom's views on the philosophy of language, as well as the conception of scientific change in general associated with it.

Keywords versions of a theory, variational principles, classical mechanics, origin of scientific change.

1. RATIONAL RECONSTRUCTION AND UNDERSTANDING

There are two main ways of talking about scientific theories in the history of science: either they are taken as opposing blocks on either side of a scientific revolution, or they are taken as evolving by being confronted with new data. When revolutionary change is considered, the pre-revolutionary theory is usually presented in a rather simplified, or even schematized fashion. For example, Newtonian mechanics is often oversimplified when describing the quantum or the relativity revolutions. It is detached from its own history and only seen from the point of view of the theory, or theories, succeeding it. As for intra-theoretical change, it is usually viewed as driven by new empirical findings.

In this paper, I focus on a mode of scientific change that does not fall into either category. I claim that some theories evolve on their own, not in response to new empirical data. I suggest that this mode of scientific change may shed light on scientific change in general since most cases of intra-theoretical change are likely to involve their share of internal evolution not determined by new empirical findings. It is also plausible that parts of revolutionary change are governed by strictly internal determinations. My aim

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 87–101.

is to determine the kinds of problems specific to this mode of scientific change and to propose some methodological tools for solving these problems.

In order to understand what internal intra-theoretical change is and what questions it raises about scientific change in general, it is first necessary to break the theoretical blocks, so to speak, that stand facing one another as a consequence of a scientific revolution. This implies distinguishing different versions of a scientific theory, namely taking into account that its meaning and potential for evolution are themselves subject to change.

A version of a scientific theory can be primarily defined as the way a user (or a group of users) of the theory understands it. A user of a scientific theory may be a scientist working at improving it, and/or a teacher transmitting it to students. This first, crude definition may seem at odds with the usual conceptions of scientific theories, which do not appeal to the way people understand them, but rather to features independent of scientists' minds. In focusing on the first-person use of scientific theories, I am not taking sides in the debate between syntactic and semantic conceptions. I take theories to be sets of *interpreted* sentences and sophisticated tools of representation (examples of which are Humphreys' templates[1]) as they are understood and used by scientists.

However, insisting on versions of a scientific theory that are too small in scale has many drawbacks from a methodological point of view. Is there no intermediate path between the multitude of individual, but intractable versions of, e.g., Newtonian mechanics that one finds when one pays attention to intra-theoretical scientific change and the simple, but unrealistic, versions that serve to mark the contrast with the theory of relativity or with quantum mechanics? Some realistic versions – i.e., ones faithful to the (abstract) objects that the scientists are handling at a time – are needed in order not to be too schematic in the historical description. But the number of versions involved seems to dilute the very notion of scientific change, and, moreover, to confuse the notion of *intra*-theoretical change, i.e., change within the *same* theory.

To escape the dilemma of methodological intractability versus oversimplification, I propose a micro-analysis of intra-theoretical change in terms of making the content of some version of a scientific theory more explicit. In Sect. 2, I try to bring out various aspects of intra-theoretical change by describing parts of the history of classical mechanics. In Sect. 3, I try to answer questions emerging from my example, and in Sect. 4, I present a new analysis of scientific change, relying on some insights of Robert Brandom's philosophy of language.

2. A SAME THEORY, DIFFERENT VERSIONS

As is well-known, there is a sense of "sameness" in which Newtonian, or classical mechanics was not *the same* in 1687 and in 1900. There is also a sense of "sameness" in which it is meaningful to speak of the persistence of classical mechanics as "one and the same" theory over three centuries time. On what grounds can both these claims be true?

In order to answer this question, as my main example, I shall focus on some parts of the history of variational principles. Classical mechanics can be presented, expressed,

[1] Cf. Humphreys (2004).

or formulated in two guises: with the help of equations of motion, or with the help of variational principles. Variational principles are, in some sense, which is worth examining, equivalent to equations of motion. Summarizing a view shared by many, the physicist Max von Laue, writes that one can even say that variational principles contain equations of motion, a claim we shall have to come back to below (Laue, 1950). An initial, coarse-grained analysis of this equivalence is that variational principles and equations of motion have the same empirical consequences, i.e., that the only difference between these two formulations is *formal* or, more precisely, that there is a chain of mathematical deductions leading from one to the other. Comparing Newton's initial formulation, involving equations of motion and Jacobi's formulation, some 150 years later, in terms of a variational principle, we thus have a sense in which they express the *same* theory. However, upon a finer-grained, diachronic analysis this sense begins to elude our grasp.

The first formulation of (part of) mechanics by means of a variational principle – as we now call it – is due to Maupertuis in 1746 in a paper called "Les lois du mouvement et du repos déduites d'un principe métaphysique" (Laws of motion and rest deduced from a metaphysical principle).[2] Maupertuis had first introduced the principle of least action in optics in 1744.[3] One of his aims in 1746 was to provide mechanics with a theological foundation by affirming that nature acts in a manner that makes some quantity a minimum. Through experimentation, he found that this quantity depends on mass, velocity, and distance. He called the product of the three factors "action" and accordingly expressed a "principle of the least quantity of action": "La Nature, dans la production de ses effets, agit toujours par les moyens les plus simples. (…) Lorsqu'il arrive quelque changement dans la Nature, la quantité d'action nécessaire pour ce changement est la plus petite qu'il soit possible." ("In producing her effects, Nature always acts using the simplest means. (…) Whenever some change occurs in Nature, the quantity of action necessary for this change is the least possible".). Maupertuis claimed to derive the definition of quantity of action from metaphysical considerations about God's aims and to have obtained a universal principle of physics. However, he established the validity of the principle only in mechanics, for systems of masses subject to central forces.

Even if the definition of action as the product *mvs* is by no means satisfactory, because distance varies with time, and even if Maupertuis' derivations are full of errors, this is an interesting example of intra-theoretical scientific change. A new concept (action) is introduced, and an effort is made to give the principles of mechanics sounder foundations. (Many physicists of the time thought that Newtonian mechanics lacked a secure foundation and looked for it in God. The concept of force was particularly suspect.) Both introduction of new concepts and attempts to provide a better justification for the principles of a theory can be counted as ingredients of intra-theoretical change from a practical point of view. New methods for studying mechanical problems have been introduced. From an architectonic point of view, which was a main concern at the time, the place of mechanics within knowledge has changed.

2 Maupertuis (1748b).
3 Maupertuis (1748a).

In the same year Euler (1744) published a different version of a principle bearing the same name, the "principle of least action", and partly the same meaning, but associated with wholly different motivations. Euler's paper is called "Methodus inveniendi lineas curvas, maximi minimive proprietate gaudentes sive solutio problematis isoperimetri latissimo sensu accepti" ("Method for inventing curve lines with a property of maximum of minimum which are solutions to problems of isoperimeter taken in the broadest sense"). This title shows that the methods favored by Euler are mainly geometrical. In an appendix, Euler shows how to solve the problem of the motion of a material point subjected to central forces. He explicitly restricts himself to conservative systems (in modern terms).

Euler was deeply convinced that there are two equivalent modes of explanation of physical phenomena, corresponding to two types of causation, namely by efficient and by final causes. Both modes of explanation contribute to the building of a grand house of physics (Pegny, 2005).[4] In 1744, he tried to show that the dynamics of certain conservative systems can be explained and predicted by what he conceived of as final causes, expressed by the principle of least action.

In the 1744 paper, the principle assumed the status of an exact mathematical statement stating that when a particle travels between two fixed points, it takes that part for which $\int v ds$ is a minimum. More precisely, the principle can be stated as follows:

$$\delta\int_P^Q v ds = 0$$

where "P" and "Q" refer to the initial and final points of the path, and δ refers to the variation of the integral under the restriction that energy is fixed. In a paper published in 1751 ("Harmonie entre les principes généraux de repos et de mouvement Euler (1751)"[5]), Euler's purpose is to lay down the principle of least action as a universal principle of dynamics. He managed to generalize it to continuum mechanics.

Euler's version of the principle of least action displays new aspects of intra-theoretical change. Here, there is no sign of any theological concern, only mathematical precision and demonstration. Seen from the point of view of the overall meaning of mechanics, this involves a change in perspective, at least in comparison with the concerns Maupertuis had. This change in perspective – from theology to mathematics – induces a change in content. From Euler's point of view, part of the content of mechanics is mathematical and has to do with proving theorems and generalizing their field of application with the help of purely formal tools. This does not mean that from Euler on, the development of mechanics was exclusively achieved along mathematical lines; but it suggests that intra-theoretical scientific change may occur in various, intertwined intellectual dimensions.

Continuing within the history of mechanics, we come across another commonly discussed aspect of inter-theoretical change, namely generalization. Lagrange proved

[4] See M. Pégny's Master thesis for more details about Euler's conviction (Pegny, 2005).
[5] Euler (1751).

the validity of the principle of least action for cases much more general than those Euler studied. In a paper called "Application de la méthode exposée dans le mémoire précédent à la solution de différents problèmes de dynamique" (Application of the method presented in the previous paper [namely the calculus of variations] to the solution of various problems of dynamics),[6] he indeed treated a system of mutually interacting particles, the forces between them deriving from a potential. It would be too long to list all the problems studied by Lagrange. Let us just mention: the motion of a body attracted by any number of fixed centers by forces proportional to any functions of the distances; the motion of a system of some bodies, each subjected to any number of central forces and acting on one another through any number of mutual attractive forces; the orbits of three bodies exerting mutually attractive forces on each other relative to a fourth body considered as fixed, etc.

After having defined or re-defined the action of each particle as $\int vds$, he stated the principle of least action as affirming that the system of interacting particles moves from one configuration to another in such a way as to make the total action, i.e., the sum of the actions of the individual particles, stationary as compared to adjacent virtual motions between the same initial and final configurations and having the same energy as the actual motion. Formulated in words: 'Soient tant de corps qu'on voudra M, M', M'', …, qui agissent les uns sur les autres d'une manière quelconque, et qui soient de plus, si l'on veut, animés par des forces centrales proportionnelles à des fonctions quelconques des distances; que s, s', s'', …, dénotent les espaces parcourus par ces corps dans le temps t, et que u, u', u'', …, soient leurs vitesses à la fin de ce temps; la formule $M\int uds + M'\int u'ds'$ +, etc. sera toujours un maximum ou un minimum'. In symbols:

$$\delta_{E\ const.}\left(M\int_A^B uds + M'\int_A^B u'ds' + \ldots\right) = 0$$

where A is the initial, and B the final configuration. Lagrange pursued this enterprise of generalization, and in the *Mécanique analytique*, he extended the principle of least action to entirely general dynamical systems, even dissipative ones. However, his mechanics remains a mechanics of forces: he considerably minimizes the physical novelty of the variational calculus.

How are we to characterize Lagrange's pursuit of an entirely general formulation of the least action principle? Is it an instance of purely mathematical change?[7] In this case, the written expression of the principle has to be conceived of as an uninterpreted formula, void of any physical meaning. On the contrary, Lagrange's work may also be seen as a systematic investigation into the consequences of the laws of mechanics, i.e., an investigation into the meaning of these laws, rather than the computation of purely mathematical implications. Lagrange himself finds it worth specifying that, for him, the principle of least action, as well as the principle of the conservation of energy, are not metaphysical postulates, but simple and general consequences of the laws of

[6] Lagrange (1760).
[7] Cf. Kitcher (1983) for an analysis of mathematical change.

mechanics. This claim responds to the worries induced by Maupertuis' interpretation of the principle of least action and emphasizes the formal reading of the principle. Let us quote Lagrange, in the "Avertissement de la première édition" of the *Mécanique analytique* (1853–1855):

> On a déjà plusieurs Traités de Méchanique, mais le plan de celui-ci est entièrement neuf. Je me suis proposé de réduire la théorie de cette Science, et l'art de résoudre les problèmes qui s'y rapportent, à des formules générales, dont le simple développement donne toutes les équations nécessaires pour la solution de chaque problème. (…)
>
> On ne trouvera point de Figures dans cet Ouvrage. Les méthodes que j'y expose ne demandent ni constructions, ni raisonnements géométriques ou méchaniques, mais seulement des opérations algébriques, assujetties à une marche régulière et uniforme. Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche, et me sauront gré d'en avoir ainsi étendu le domaine.
>
> (We already have several Treatises of Mechanics, but the conception behind this one is entirely new. I assigned myself the task of reducing the theory of this Science and the art of solving the problems relative to it to general formulae, the simple development of which gives all the equations necessary for the solution of each problem. (…)
>
> No Figures are found in this Work. The methods I present in it require neither geometrical constructions nor geometrical or mechanical reasoning, but only algebraic operations subjected to a regular, uniform sequence of steps. Those who like Analysis will be pleased to see Mechanics become a new branch of it and will be grateful to me for having thus extended its domain).

For Lagrange, the formulation of mechanics with the help of variational methods is, thus, no more than a pure enterprise of mathematical rewriting. No new physical meaning is added. The main benefit of such a formal rewriting is that it allows for a more efficient mastery of the problems.

Before the publication of Hamilton's work in the 1830s (Hamilton, 1834–1835), the general opinion of physicists about the principle of least action was that it sounded like an interesting item of physical theorizing, but could not lead to substantial developments in mechanics. Let us quote S. D. Poisson's well-known judgement of it: "Le principe de moindre action est une règle, aujourd'hui inutile, pour construire les équations différentielles du mouvement" (The principle of least action is a rule, useless today, for constructing the differential equations of motion). The reason for this opinion is that the principle suffers from an important limitation: it only applies to virtual paths having the same energy as the actual path, which makes it difficult to apply, because the energy of a system is often difficult to ascertain. Hamilton's achievement was to remove this restriction concerning the energy of the virtual paths. His new formulation of the principle of least action is as follows:

> [A] system moves from one configuration to another in such a way that the variation of the integral $\int Ldt$ between the path taken and a neighboring virtual path having the same initial and final spatio-temporal points is zero (here, $L = T - V$, T is the kinetic energy of the system and V its potential energy).

In symbols, Hamilton's principle is:

$$\delta \int Ldt = 0 \tag{1}$$

This allows to him establish, or better, to disclose, a direct connection between variational principles and equations of motion. The conditions for the integral $\int L dt$ to be stationary, namely:

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial f}{\partial\left(\dfrac{dx}{dt}\right)}\right) - \dfrac{\partial f}{\partial x} = 0 \\ \dfrac{d}{dt}\left(\dfrac{\partial f}{\partial\left(\dfrac{dy}{dt}\right)}\right) - \dfrac{\partial f}{\partial y} = 0 \end{cases} \tag{2}$$

with $f = f\left(x, y, \ldots, \dfrac{dx}{dt}, \dfrac{dy}{dt}, \ldots, t\right)$, are precisely Lagrange's equations of motion for the system:

$$\frac{d}{dt}\left(\frac{\partial L}{\dot{q}_r}\right) - \frac{\partial L}{\partial r} = 0 \tag{3}$$

As Hamilton's principle is independent of any coordinate system, the relation between (1) and (3) implies a new and important generalization of the laws of mechanics.

The fact that the principle of least action is verified just in the case that Lagrange's equations (3) hold shows how deeply variational principles and equations of motion are connected. Such an intimate connection sheds new light on what was considered just a nice consequence of the laws of motion. Hamilton's principle, so to speak, "says" something about mechanics that was unknown before Hamilton's work. Owing to it, the conceptual structure of mechanics looks different. A conceptual change of this importance is undoubtedly a major event in the history of a theory as old as mechanics. It reaches far beyond the generalization provided by the fact that Hamilton's principle is independent of any system of coordinates. Yourgrau and Mandelstam (1968), for instance, qualify Hamilton's achievement as a "fundamental advance in the theory of mechanics". After him, they say, the meaning, aspect and consequences of the theory have changed.

More precisely, Hamilton re-defined the action of a system as a function of the initial point of the path taken, the final point, and the energy. These ($2n + 1$) quantities are necessary and sufficient to determine the path, and hence the action, uniquely. Hamilton's "law of varying action" may be formulated as follows:

$$\delta S = 0 \tag{4}$$

with $S = S\left(q_{xa}, q_{xf}, E\right) = \left(\int_{q_{xa}}^{q_{xf}} \sum_s p_s q_s\right)_{H=E}$

Thus (4) is equivalent to

$$\delta\left(\sum_i m_i \int v_i ds_i\right) = \sum_i \sum_{x,y,z} m_i \dot{x}_{if} \delta_f x_i - m_i \dot{x}_{ia} \delta_a x_i + \int \delta E dt = 0 \tag{5}$$

and to

$$\delta \int \sum_r p_r dq_r = \int \left[\sum_r d\left(\frac{\partial L}{\partial \dot{q}_r} \delta q_r\right) + \delta H dt\right] = \sum_r p_{rf} \delta_f q_r - p_{ra} \delta_a q_r - \left(H_f \delta_f t - H_a \delta_a t\right) \tag{6}$$

Hamilton thus established that everything that was known at the time about the mathematical properties of motion is deducible from his "law of varying action". Such a unification means that the mathematical and conceptual structure of mechanics is much richer than physicists and mathematicians had usually thought it to be, as Hamilton's words cited below indicate. The "law of varying action" captures what is at the root of the various equations used to study the motion of bodies and of systems of bodies.

It is worth quoting Hamilton's own appraisal of all this:

> Yet for not having formed the conception of the action as a function of this kind, the consequences that have been deduced from the formula (5) for the variation of that definite integral appear to have escaped the notice of Lagrange. (…) For although Lagrange and others, in treating of the motion of a system have shown that the variation of this definite integral vanishes when the extreme coordinates and the constant H are given, they appear to have deduced from this result only the well known law of least action. (…) But when this well known law of least, or as it might better be called, of stationary action, is applied to the determination of the actual motion of a system, it serves only to form, by the rules of the calculus of variations, the differential equations of motion of the second order, which can always be otherwise found. It seems, therefore, to be with reason that Lagrange, Laplace and Poisson have spoken lightly of the utility of this principle in the present state of dynamics. A different estimate, perhaps, will be formed of that other principle which has been introduced in the present paper, under the name of the law of varying action, in which we pass from an actual motion to another motion dynamically possible, by varying the extreme positions of the system, and (in general) the quantity H, and which serves to express, by means of a single function, not the mere differential equations of motion, but their intermediate and their final integrals.

Hamilton insisted that before his work, the principle of least action was merely redundant with respect to the usual formulation of mechanics in terms of differential equations. His achievement lies in discovering how to derive *new* results from his "law of varying action". What could otherwise have seemed a purely formal game within variational calculus definitely proved fruitful.

The history of mechanics did not stop with Hamilton. It could be extended by presenting the new generalization of Hamiltonian mechanics provided by symplectic geometry, which is today taken as making the conceptual core of mechanics explicit. Moreover, one could also write the story of how the appraisal of Hamilton's achievement has changed. From our mathematically more sophisticated point of view, we tend to emphasize the importance of Hamilton's formulation's being independent of any coordinate system. However, we now have enough examples of varieties of internal intra-theoretical change. In Sect. 3, I shall try to classify and analyze them.

3. THE VARIETIES OF INTRA-THEORETICAL CHANGE

The above example drawn from the development of mechanics raises four main questions:

(*i*) In what sense was mechanics a unique, evolving theory from 1687 to 1840?
(*ii*) Was intra-theoretical change in mechanics purely formal?
(*iii*) Is the mainspring of theoretical change mathematica or physical?
(*iv*) Was the entire development of mechanics already implicitly present in Newton's *Principia mathematica*? (1687)

(i) As already mentioned, the answer to the first question depends on the degree of precision and purpose of the historical analysis. When classical mechanics is taken – from a post-relativity theory point of view – as the science of motion for velocities much less than the velocity of light, its formulation in terms of equations of motion, and that in terms of variational principles, are best considered to be equivalent, because they have the same empirical consequences and allow for the same empirical predictions. However, when looking diachronically, step by step, we can see that the domain of the theory expands: namely, ever more general dynamical systems are taken into account. All the particularizing hypotheses, e.g., about the material constitution of the systems, are progressively eliminated. The predictive power of the theory is greatly enhanced in this process. So Hamiltonian mechanics is not the same as Newtonian mechanics after all. The concept of force has been excluded, and a fully universal concept of physics has been used with its full power, the concept of energy. The difference is presumably not a difference of know-how, namely concerning the computational power of physicists. Of course, because of the availability of more sophisticated mathematical tools, mathematicians and physicists were much better at computing mechanical results at the end of nineteenth century than they were at the end of eighteenth century, but the scope and meaning of mechanics have also changed, and have so along the lines of generalization and unification.
(ii) If one acknowledges that mechanics has changed from Newton to Hamilton and Jacobi, then one next has to qualify this change. Some might be willing to call it purely formal in nature, i.e., to say that what happened was that the differential equations of mechanics had been better understood qua mathematical objects (vs. qua representations of empirical phenomena). In this view, the change is "formal" in the sense that it consists in mathematical elucidation, explication, or development. At the basis of this analysis lies a distinction between formal and empirical. However, was the development of mechanics devoid of any empirical content? It seems that when a theory applies to larger domains of phenomena, its empirical content is modified, which also happens when new relations are established among domains formerly considered to be separate. In those cases, our understanding of phenomena is improved, which is a good indicator of scientific change.
(iii) Euler's and Lagrange's works seem to suggest that mathematical investigation is the only mainspring of the development of mechanics. However, the distinction

> between what falls within mathematics and what is a matter of physics is not easy to draw. The various ways of solving differential equations are disclosed by mathematical analysis. Mathematical investigation drives theoretical change in mechanics only when it is given a physical interpretation, for instance, when it allows for a generalization of the theory. Now the boundary between "mathematical analysis" and "physical interpretation" is porous, especially when investigated from a historical point of view. When twentieth-century mathematicians used mathematical investigation to disclose the conceptual structure lying at the root of mechanics, they very likely established an important property of mechanics qua physical theory. In other words, mathematical investigation is twofold: on the one hand, it has formal, as opposed to empirical, implications; on the other, it can make clear what mechanics qua physical theory is about. To sum up, question (iii) is likely to remain open until we possess a better understanding of the relationships between mathematical and physical knowledge.

Dealing with questions (*i*)–(*iii*), which emerged from the investigations into the history of mechanics outlined above, clearly shows that scholars are not well-equipped to study intra-theoretical change. The usual distinctions, between the formal and the empirical, or between form and content, or between mathematics and physics are not fine-grained enough to make us understand the nature of intra-theoretical change. This should drive us to turn to an alternative way of analyzing intra-theoretical change. The question to guide this new analysis is question (*iv*): Was all the development of mechanics already implicitly present in Newton's *Principia*? In Sect. 4, I show that answering that question can yield new insights into scientific change, be it intra- or inter-theoretical.

4. MAKING SCIENTIFIC CONTENT EXPLICIT

Even if generalization and the discovery of relationships between formerly distinct domains are not "purely formal" developments, it is a fact that the evolution of mechanics that we have considered so far is purely internal to the theory itself. In other words, it is not due to the need to account for new empirical facts. As such, it may be situated at one end of a spectrum describing what drives scientific change. At the other end of the spectrum lie cases where change is due to unexpected empirical discoveries. The spectrum metaphor is meant to point out that usual cases of scientific change are mixed, namely both internally and externally driven within varying proportions. The part of the history of mechanics presented in Sect. 2 is one of the few examples that we know of purely internal scientific change.

The above claim about the forces driving scientific change runs counter to Kuhn's thesis about the development of new formalisms as presented in "Second thoughts about paradigms":

> [S]pecial formalisms are regularly accepted as plausible or rejected as implausible in advance of experiment. With remarkable frequency, furthermore, the community's judgments prove to be correct. Designing a special formalism, a new version of the

> formalization, cannot therefore be quite like inventing a new theory. Among other things, the former can be taught as theory invention cannot.[8]

Kuhn's thesis, which emerges from the explicit opposition between "special formalisms" and "theory", is that the invention of a new theory is necessarily driven by empirical discovery, in contraposition to the designing of new formalisms. I think that Hamilton's achievement is an important counter-example to this thesis because it manifests genuine theoretical novelty: his work can by no means be described as "designing a special formalism", or "a new version of the formalization". Taken as a whole, throughout their history, variational principles built up a new genuine theory of mechanics, even though it is not new in the sense usually associated with the word "invention". The psychological category of invention is probably not the only relevant one when dealing with theoretical change, since it conceals an important distinction between two types of theoretical novelty, the former propelled by empirical discoveries and the latter a result of conceptual investigation. Whereas Kuhn is only concerned with the former, I think that the latter is of great importance in the development of science, in the study of internally as well as externally driven scientific change.

Internal change may be described as elucidation of a formerly present content. As this notion is rather elusive, it is worth going back to the historical example of Sect. 2. The main difficulty in writing (part of) the history of mechanics was already clear from Max von Laue's quote: "Since Euler, mathematics has set down principles of variation which were equivalent to equations of motion: one can even say that they contained them". In order to avoid the conclusion that there was no empirical or physical change in mechanics through the nineteenth century, we have to find a way to understand Laue's claim that would allow for new scientific content to emerge in spite of its already being implicitly present.

I suggest that we can find a useful tool for this task in the notion of "version of a scientific theory" used above. A version of a scientific theory is not only a product, namely a set of statements, concepts, problems, and ways to solve them; it is also the production of a mind (or of minds) at work. As was said above, all physicists using mechanics virtually have their own versions of it, depending on how they learned it, what books they read, what problems interested them – and on their inferential capacities. Taking this idiosyncrasy seriously amounts to looking for the *reasons* physicists have for using and sometimes for developing one version of a theory or another. This in turn amounts to looking for their reasons for *changing* the theory, sometimes in minor ways, sometimes more significantly. At the bottom of such a capacity for developing a personalized version of a scientific theory lies a capacity for understanding (some of) previous versions of it, and in particular the concepts they involve.

This suggestion may be contrasted with Kuhn's claims about the "formal, or readily formalizable, components"[9] of a disciplinary matrix: He claims that the symbolic generalizations of scientific theories, which are the main examples of the "formal components" of a matrix, are commonly applied as a matter of routine, "without felt

[8] Kuhn (1977, p. 301).
[9] Kuhn (1977, p. 297).

need for special justification".[10] Symbolic generalizations are often sketches of full generalizations, the detailed expression of which varies according to the application considered. This is why, in Kuhn's view, their content remains unquestioned during periods of normal science. However, the development of mechanics shows that even generalization sketches are investigated *qua* meaningful generalizations, leading to new methods which in turn can be used as routines – this was Lagrange's aim. The very meaning of the symbolic generalizations may be a serious domain of scientific investigation within a given disciplinary matrix, contrary to Kuhn's slightly oversimplified conception. Nevertheless, Kuhn's claim that the Lagrangian or Hamiltonian formulations of mechanics are good examples of the usual evolution of mathematized disciplines, namely of a tendency to develop sketches of laws to be adapted to each particular problem,[11] points to an important aspect of intra-theoretical change.

Looking at what happens from the point of view of the users and producers of the successive versions of mechanics is for me a useful and fruitful way to build up a well-documented, external point of view, namely the point of view of the historian. Historians of science usually contemplate the products of the practice of using and developing mechanics, i.e., its versions themselves. In order to give faithful and meaningful descriptions of these versions, they have to turn to the special form of intellectual practice that characterizes scientific theorizing.

The first requirement for a good practice of mechanics is to understand the concepts used by one's predecessors, namely, following Robert Brandom, "to master their roles in the relevant inferences".[12] This amounts to "knowing what else one would be committing oneself to by applying the concept, what would entitle one to do so, and what would preclude from such entitlement".[13] What is involved, according to Brandom, in concept use? He writes that

> [I]n making a claim, one is implicitly endorsing a set of inferences, which articulate its conceptual content. Implicitly endorsing those inferences is a sort of doing. Understanding the conceptual content to which one has committed oneself is a kind of practical mastery: a bit of know-how that consists in being able to discriminate what does and what does not follow from the claim, what would be evidence for and against it.[14]

This may lead to a good explanation of the fact that, although the further developments of mechanics are already implicit in Euler's version, his successors have nevertheless made genuinely new contributions to it. They have worked at a better understanding of the concept of action and of the claim that, in certain cases, the quantity of action is minimized. In doing so, they have exercised their capacities of discriminating what are true, or at least plausible, consequences of this claim. Brandom's scheme points to the fact that scientific theorizing, even when it is not pressed by new empirical facts, is a demanding cognitive activity, requiring abilities traditionally associated with invention.

[10] Ibid., p. 298.
[11] Ibid., note 10, p. 300.
[12] Brandom (2000, p. 11).
[13] Ibid.
[14] Ibid., p. 19.

Resorting to the way scientists understand theories in order to analyze scientific change is rather unusual.[15] However, it is a promising way to solve traditional problems in the history of science, for instance the question as to whether the development of mechanics as it actually happened was inevitable or not. From a point of view that does not take the scientists' cognitive capacities into account, it seems that, to the extent that later versions of mechanics are already contained in Euler's formulation, the development of mechanics was necessary and inevitable. However, emphasizing the role of concrete human beings in scientific change helps understand that nothing was inevitable in it. It may have happened that either nobody would have been interested in the conceptual development of mechanics, or that someone would have developed symplectic geometry a 100 years before it was actually developed, or.… Emphasizing what scientists actually do with theories is a fruitful way to make sense of the contingency of their history. The topic of scientific theories' historical contingency is however much larger than the former sentence suggests. For instance, the question whether the theory describing and explaining motions of medium-size objects at speeds much smaller than the speed of light could have happened to become incompatible with Newton's original principles cannot be answered by only taking the scientists' cognitive capacities into account.

Applying Brandom's views on the example presented in Sect. 2 leads us to see intra-theoretical change, in the case of mechanics, as a process of explication, or of expression of a content that was already there. I insist that this does not amount to minimizing the novelty involved in this change. In the process of expressing the content of Newtonian mechanics, i.e., of making explicit what was implicit in Newton's (or Maupertuis', or Euler's) formulation, physicists have conceptualized their subject matter, in the sense that they have applied concepts in a novel and more self-conscious way. This new way of using the concepts of mechanics consists in a better appraisal of their roles in reasoning. For instance, the successive definitions of the concept of action show how Lagrange and Hamilton explored the implications of considering this quantity as a minimum in terms of generality of the theory and of its facility of application.

Let us now turn to the tricky questions presented in Sect. 3. When the mainspring of theoretical change is internal, and partly mathematical, there is a sense in which later versions of the theory may be said to have the same content as earlier ones, in spite of the unsatisfactory character of that phraseology. However, when one seriously takes into account how rich the implicit content of any claim is (except analytic ones), the "sameness of content" claim can be made to cohere with the emergence of genuine conceptual novelty.

> How can the implicit content of any claim be made explicit? According to Brandom, "an explicit claim has implicit in it:
>
> (1) properties governing inferential moves to and from the commitments to the claimable content in question;

[15] At least in the papers and books analysing the products of scientists' activity. A different tradition that has been rendered illustrious by Koyré, Kuhn, Hanson, and more recently by Léna Soler, focuses more generally of scientific *thinking*, of which understanding is a small part.

(2) the other claims that are inferential consequences of the first one, according to the properties mentioned in (1);
(3) the conceptual content of the claim, which is articulated by the inferences in (1).[16]

The development of mechanics can be seen as the active articulation of the conceptual content of Newton's three principles by means of a conceptualization of the quantity which was named "action".

Finally, in Brandom's analysis of the use of concepts in communication, we can find a way to give a precise meaning to the notion of a version of a scientific theory. Any specification of a propositional content, he says, must be made from the perspective of some set of commitments. From this perspective, "one wants to say that the *correct* inferential role of a concept is determined by the collateral claims that are true. Just so; that is what *each* interlocutor wants to say: each has a slightly different perspective from which to evaluate the inferential properties".[17] As every physicist accepts a slightly different set of propositions from his or her predecessors, the inferences the physicist draws from their versions of mechanics, for instance, differ from the inferences they have drawn, according to the explications the physicist gives of the relevant concepts. When able to exercise inferential capacity with maximal efficiency, the version that the physicist develops involves a new explicit content, that at the same time shares the same content as that of its predecessors.

5. CONCLUSION

I have tried to give an account of intra-theoretical change by using Brandom's notion of "making a claim explicit", which focuses on the inferences allowed by the claims one accepts. I have insisted on the fact that, usually, scientific change is both externally and internally driven, and that the example presented in Sect. 2 is exceptional. However, I claim that the internally driven part of scientific change has to be taken into account, which implies closely investigating the inferential moves undertaken by the scientists involved. Since scientific communities are composed of individual scientists with "slightly different perspectives", engaged in communication and subtle inferential and linguistic games, the individualistic methodology does not run counter to Kuhn's emphasis on communities as much as it seems prima facie to be the case. The social dimension of scientific practice is fully accounted for by the essentially social nature of language and inferential practice.

Acknowledgments

Many thanks to Léna Soler who has been so patient, and to an anonymous referee who helped me to make my claims more precise. I also want to warmly thank Maël Pegny, who did an excellent job for his master thesis, and Marion Vorms, who so efficiently

[16] Ibid., p. 19.
[17] Ibid., p. 183.

took on my own questions. Finally, I thank Alexandre Guay for his support in this enterprise.

BIBLIOGRAPHY

Brandom, R. (2000) *Articulated Reasons. An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Euler, L. (1744) Methodus inveniendi lineas curvas, maximi minimive proprietate gaudentes sive solutio problematis isoperimetri latissimo sensu accepti. In *Leonhardi Euleri Opera omnia. Seris Prima. Opera mathematica*. Volumen Vicesimum quartum. Edidit Constantin Carathéodory. Orell Füssli Turici. Berne, 1952.

Euler, L. (1751) Harmonie entre les principes généraux de repos et de mouvement de M. de Maupertuis. *Histoire de l'Académie Royale des Sciences de Berlin pour l'année 1751*.

Hamilton, W. R. (1834–1835) *On a General Method in Dynamics*. *Philosophical Transactions of the Royal Society*, Vols. 124 and 125.

Humphreys, P. (2004) *Extending Ourselves*. New York: Oxford University Press.

Kitcher, P. (1983) *The Nature of Mathematical Knowledge*. New York: Oxford University Press.

Kuhn, T. (1977) *The Essential Tension. Selected Studies in Scientific Tradition and Change*. Chicago, IL/London: University of Chicago Press.

Lagrange, J.-L. (1853–1855) *Mecanique Analytique*, 3e édition, revue, corrigée et annotée par M.-J. Bertrand. Paris: Mallet-Bachelier.

Lagrange, J.-L. (1760) Application de la méthode exposée dans le mémoire précédent à la solution de différents problèmes de dynamique. *Miscellaea Taurinensis*, t. 2, 1760–1761.

von Laue, M. (1950) *History of physics*. O. Ralph (trans.). New York: Academic.

Moreau de Maupertuis, P.-L. (1748a) Les lois du mouvement et du repos déduites d'un principe métaphysique, *Histoire de l'Académie Royale des Sciences de Paris pour l'année 174*6. Paris.

Moreau de Maupertuis, P.-L. (1748b) Accord de différentes lois de la nature qui avaient jusque ici paru incompatible. *Histoire de l'Académie Royale des Sciences de Paris pour l'année 1744*. Paris.

Newton, L. (1687) *Principia mathematica philosophiae naturalis*.

Pegny, M. (2005) *Le principe de moindre action de Maupertuis à Lagrange*. Master thesis, Université Paris X – Nanterre.

Yourgrau W. and Mandelstam, S. (1968) *Variational Principles in Dynamics and Quantum Theory*. 3rd ed. Philadelphia, PA: Saunders.

COMMENTARY ON "FROM ONE VERSION TO THE OTHER: INTRA-THEORETICAL CHANGE", BY ANOUK BARBEROUSSE

IGOR LY

In this commentary, I will only raise a couple of methodological questions that may help clarify the rich content of Barberousses's paper.

While taking into account an interesting feature in the evolution of physics – the process of making explicit the content and meaning of a physical theory – Barberousse tries to develop an important topic, namely the central difficulty one encounters when comparing scientific theories and considering scientific revolution: how to define the parameters of what we call "a theory." In fact it is often difficult to apply to concrete cases Kuhn's notion of "normal science" and the various features he attributes to it. Barberousse gives some examples of these difficulties in the case of "classical mechanics." I will only add one more comment. Newton writes:

> I do not feign hypotheses. For whatever is not deduced from the phenomena must be called a hypothesis; and hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy. In this experimental philosophy, propositions are deduced from the phenomena and are made general by induction. (Newton, 1999, p. xii)

D'Alembert writes:

> Nous ne l'adopterons pas non plus [d'Alembert is speaking about the principle according to which "la force accélératrice ou retardatrice est proportionnelle à l'élément de la vitesse"], avec quelques géomètres, comme de vérité purement contingente, ce qui ruinerait la certitude de la Méchanique, et la réduirait à n'être plus qu'une Science expérimentale […]. (D'Alembert, 1990, p. 943)
>
> "We shall not borrow, along with some geometers, as a merely non necessary truth, which would ruin the certainty of Mechanics, and would reduce it to be only an experimental science [...]" (my English translation)

When looking at these opposing claims, can we say that Newton and D'Alembert work within the same paradigm of mechanics?

One of Barberousse's main achievements in this paper is to show how complex it is to determine what we mean by a "same theory," and "different versions of the same theory" especially because of the changing nature of the "same theory" from the moment we try to apply it to a single scientific work. I would particularly emphasize that in the example she studies Barberousse shows how problematic and often

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 103–105.

misleading it is to distinguish between form and content when describing a physical theory.

Her adopted point of view consists precisely in taking into account the dynamics and evolving character of what we call "same theory" – in other words *intra-theoric change* – and in trying to define the sameness of a theory through its various versions by defining that character. Her idea is to create an important basis for criticizing the notions of paradigm and "normal science", and use it as the main tool to give these notions sense and legitimacy.

In my commentary I will raise two questions about Barberousses's thesis, or proposed answer. The first question concerns the status of her answer. The second one is directed at its content and meaning.

1. THE STATUS OF THE ANSWER

In what sense can we say we are dealing with various versions of the same theory? If we consider that this is Barberousse's main question we may wonder whether the proposed answer is explanatory or normative.

If it is a matter of understanding why some mechanical works are said to belong to the same theory, in this case the theory of "classical mechanics", then the answer is explanatory. It is important there to determine who says it, and in what circumstances. Without answering these questions first, we cannot answer the "why" question.

Or, if it is a way for historians to classify works, independently from whether they are said to belong or not to the same theory, then two questions arise. What philosophical or historiographic contribution is expected from such a construction? And are not we running into a methodological circle, since the reasons for adopting the criterion of sameness are conditioned by first accepting the fact that different scientific works belong to the same theory?

In other words, are the theoretical changes that Barberousse tries to characterize considered from the very beginning to be intra-theoretical, or are they expected to provide a tool to help decide the sameness of a theory? The difficulty lies in the paper's historiographic nature. The answer depends on Barberousse's conception of the nature and aim of the History of Science involved, and on the ways her considerations are related to that conception.

2. THE CONTENT AND MEANING OF BARBEROUSSE'S THESIS

The general question raised in the preceding section involves methodological counterpoints to the paper's thesis. Barberousses's thesis consists in focusing the attention on the understanding of a theory by the scientists who develop it, and in claiming that this process constitutes an important illustration of "scientific change" described by her as the process of "making explicit" the content of a given theory. It is worth noting here that Barberousse rightly claims that such a process does not imply a lack of creation or invention.

Nonetheless, the expression "making explicit" and the reference to Brandom might be questioned. As a matter of fact these suggest a "community of meaning." However if one considers Hamilton's work as an example of the principle of least action, it does not seem to be the case. As Hamilton says, the interpretation of this principle is very different in his own work and in his predecessors' work. Is it then possible for us to give a meaning to the expression "making explicit an implicit conceptual content" that would enable us to describe both Lagrange's and Hamilton's views about the principle of least action?

One could answer that these are two ways of explaining and making explicit – which is what Barberousse claims. But we could then say that Einstein's thoughts about simultaneity are also a way of understanding – making explicit – Newtonian mechanics. There arises a problem, which Barberousse herself somehow raises: regarding scientific change the emphasis on the process of "making explicit" seems to apply not only to intra-theoretic change, but also to extra-theoretic change. Is it then enough to say like Kuhn that Einstein's work involves a completely different conception of time and space from that of Newton's. It is not sure since classical scientists within the field of "classical mechanics" have also different conceptions of time and space. In other words, if we are looking for a criterion of sameness between different physical theories, isn't the "making explicit" criterion too broad?

Barberousse appears to be addressing two questions, sameness on one hand, and scientific change on the other, that cannot receive a single answer when one is looking at the process of "making explicit" a given scientific work.

In conclusion, Barberousse raises a very important historiographic point about scientific change – in what way can we say that two scientific works belong to the same theory ? – and offers a rich suggestion: scientific change is often the result of the process of deepening and making more explicit the content of a preceding theory. Regarding the first point, Barberousse shows us in her selected example the complexity of the question and underlines important problems encountered by Kuhn's conceptions. With the second point, in spite of not giving a totally convincing answer to the question of sameness, Barberousse demonstrates that it is a very interesting theme in the History of Sciences. As she shows it, scientists themselves often claim that in their own work they achieve a deeper understanding of their predecessors' ideas. Paying attention to the different forms and meanings of this scientific feature may constitute a valuable conceptual tool in the History of Sciences, and in the philosophical thoughts regarding scientific change.

BIBLIOGRAPHY

D'Alembert (1990) *Traité de dynamique*. In G. Jacques (ed.). Paris.

Newton, I. (1999) Mathematical *Principles of Natural Philosophy*, I. B. Cohen and A. Whitman (trans.). Berkeley/Los Angeles: University of California Press.

PART 4

INVESTIGATING THE CONTINUITIES OF SCIENTIFIC THEORIZING: A TASK FOR THE BAYESIAN?

MODELING HIGH-TEMPERATURE SUPERCONDUCTIVITY: CORRESPONDENCE AT BAY?

STEPHAN HARTMANN

Abstract How does a predecessor theory relate to its successor? According to Heinz Post's General Correspondence Principle, the successor theory has to account for the empirical success of its predecessor. After a critical discussion of this principle, I outline and discuss various kinds of correspondence relations that hold between successive scientific theories. I then look in some detail at a case study from contemporary physics: the various proposals for a theory of high-temperature superconductivity. The aim of this case study is to understand better the prospects and the place of a methodological principle such as the Generalized Correspondence Principle. Generalizing from the case study, I will then argue that some such principle has to be considered, at best, as *one tool* that might guide scientists in their theorizing. Finally I present a tentative account of why principles such as the Generalized Correspondence Principle work so often and why there is so much continuity in scientific theorizing

Keywords Bayesianism, constructionism, correspondence principle, modeling, scientific realism, theory change.

1. INTRODUCTION

Philosophers of science provide us with idealized accounts of science. Sadly, however, these accounts often do not work, as the endless series of discussions among philosophers shows. These discussions typically follow a common scheme: In step 1, a philosopher suggests a theory of *X* (say, explanation, theory change or what have you). In step 2, other philosophers criticise this account. Some point out internal problems of the account in question, others present a counterexample. Such a counterexample shows that the account in question does not always apply and this is taken to effectively refute it alltogether. In step 3, a new universal account is suggested and it is shown, perhaps, that it deals well with the counterexamples of the previous account. But it is soon shown to have other problems. And so on. Philosophers typically come up with universal and all-encompassing accounts, often based on or motivated by

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 109–129.

some simple and plausible principle which, at the end of the day, almost inevitably fails when confronted with the practice of science.

Scientists, on the other hand, proceed in a different way, which – if successful – combines a bottom-up (i.e., data-driven) methodology with a top-down (i.e., theory-driven) methodology. They construct models which account for a comparably small set of phenomena and do not intend for them to be universal in scope. Models are local, not global or universal, and once a models fails, its domain of applicability is (typically, though not necessarily) restricted and a new model with a (typically, though not necessarily) wider scope is put forward. Most importantly, scientists are aware of the fact that models involve idealizations and that they do not provide final answers to all questions they have about the object or system under investigation. And yet, models serve various purposes in the process of science (explanation, prediction, policy recommendations, etc.), and they do so very well, which is arguably one of the reasons why science as a whole is so successful. While some models in science are formulated in the framework of a theory ("models of a theory"), such as classical mechanics, others are constructed in the absence of a model-constraining theory ("phenomenological models"). A stock example of a model of a theory is the model of a pendulum, and Bohr's model of the atom is an example of a phenomenological model (see Frigg and Hartmann, 2006).

I hold that philosophers of science can learn much from scientists. They should be more modest and aim at constructing models, not theories that aim at getting everything right in just one shot. As in science, more general principles might be applied here or there, but their applicability has to be approached critically. Rather than arguing for this point *in abstracto*, I will present an example from the philosophy of science that indicates how my proposal works.[1]

The philosophical discussion of scientific theory change illustrates my approach very well. Leaving somewhat more fine-grained positions aside, two main accounts of scientific theory change can be identified. On the one hand, there is the traditional cumulative account according to which science progresses by adding more and more details to already existing accounts. This view, which stresses the *continuity* of scientific theorizing, met serious criticisms when confronted with episodes from the history of science. Inspired by this criticism, Kuhn, Feyerabend and others suggested an account of scientific theory change that stresses *discontinuities* in scientific theorizing.[2] Buzz words like "scientific revolution" and "incommensurability" figure prominently in this account, which gives, just like the cumulative view, a universal answer to the question of how scientific theory change works. One universal philosophical account is replaced by another universal account, but both get in trouble when confronted with cases from real science, as scientific theory change involves both continuity (or stability) and discontinuity (or instability).

There is certainly much more continuity in scientific theorizing, even across revolutions, than Kuhn and his followers made us think. And there might be even more

[1] For more on my views on modeling in philosophy of science, see Hartmann (2008).

[2] This is not to say that there is no place for continuities in Kuhn's philosophy of science. In *The Structure of Scientific Revolutions*, Kuhn (1996) devotes a whole chapter to a discussion of "normal science", which is characterized by an accumulation of solved puzzles. Scientific theory change, however, is discontinuous.

continuity in the future. In *The Social Construction of What?*, Hacking elaborates this point:

> [F]uture large-scale instability seems quite unlikely. We will witness radical developments at present unforseen. But what we have may persist, modified and built upon. The old idea that sciences are cummulative may reign once more. Between 1962 (when Kuhn published *Structure*) and the late 1980s, the problem for philosophers of science was to understand revolution. Now the problem is to understand stability. (Hacking, 1999, p. 85)

So how can the prevalent stability in science be understood philosophically? As the simple cumulative model does not work, an account is needed that stresses, besides the importance of continuities (or stability), the inevitable presence of more or less disruptive discontinuities (or instability) in scientific theorizing. Such an account is hard to come by and I will not present a fully worked out version of it in this contribution. Instead, I will point out a reasonable way for how one should proceed to arrive eventually at such an account. Following the scientist's strategy, I propose to combine a top-down with a bottom-up strategy and will proceed in two steps: First, I will examine examples from real science to obtain an account of the various ways in which scientific theories relate to their predecessors. This step follows a bottom-up strategy. Second, I will try to understand philosophically the prevalent continuitiy (or stability) in scientific theorizing, as pointed out by Hacking and as suggested by my case studies. This step proceeds in top-down fashion, as I will relate the findings of the case study to a more general, though sufficiently flexible, philosophical theory.

To set the scene, I will start with a critical discussion of Heinz Post's Generalized Correspondence Principle (Sect. 2). I will then outline and discuss various kinds of correspondence relations that hold between successive scientific theories (Sect. 3). Section 4 then looks in some detail at a case study from contemporary physics: the various proposals for a theory of high-temperature superconductivity. The aim of this case study is to understand better the prospects and the place of a methodological principle such as the Generalized Correspondence Principle. Generalizing from the case study, I will then argue that some such principle has to be considered, at best, as *one tool* that might guide scientists in their theorizing (Sect. 5). Finally, in Sect. 6, I present a tentative account of why principles such as the Generalized Correspondence Principle work so often and why there is so much continuity in scientific theorizing.

2. POST'S GENERAL CORRESPONDENCE PRINCIPLE

Post's General Correspondence Principle is a generalization of the quantum mechanical correspondence principle.[3] This principle played a crucial role for Niels Bohr and others in the process of constructing the new quantum mechanics in the 1920s. It was expected that quantum mechanics would account, within certain limits, for the well-confirmed phenomena of classical physics. The quantum mechanical correspondence principle is however somewhat more complicated, as Radder (1991) has shown.

[3] This section and the next draw on material published in Hartmann (2002).

The latter consists of various interrelated parts which I will not discuss here. In a first attempt, Post gives the following characterization of 'his' General Correspondence Principle:

> Roughly speaking, this is the requirement that any acceptable new theory *L* should account for its predecessor *S* by 'degenerating' into that theory under those conditions under which *S* has been well confirmed by tests. (Post, 1971, p. 228)

The General Correspondence Principle is claimed to be valid even across scientific revolutions. It presupposes that the predecessor theory *S* and the successor theory *L* "refer (in their statements) to at least some events or facts which are identifiably the same" (Post, 1971, p. 220), or, to phrase it differently, that *S* and *L* share a common set of phenomena. The domain of *L* is assumed to be larger than the domain of *S* and the account given by *L* will usually be more precise (or at least not less precise) than the account of the phenomena given by *S*. A typical example is the relation between classical mechanics and the special theory of relativity. The latter theory also correctly describes particles with a velocity close to the speed of light and provides a more accurate account at low velocities than the former.

Post goes on to discuss several possible relations between *S* and *L* that range from a complete reduction (which seems hardly ever to occur in science) to approximate or inconsistent correspondence, but without explanatory losses (such as the just mentioned relation between classical mechanics and the special theory of relativity). Other possible relations between *S* and *L* which exhibit losses would count as evidence against the General Correspondence Principle; Post holds that these relations never occured in the history of science of the last 300 years – apart from one noteworthy exception that will be discussed below.

One of Post's favorite examples to support the General Correspondence Principle is the periodic system, which survived the quantum mechanical revolution.[4] Post explains:

> The periodic system is the basis of inorganic chemistry. This pattern was not changed when the whole of chemistry was reduced to physics, nor do scientists ever expect to see an explanation in the realm of chemistry which destroys this pattern. The chemical atom is no longer strictly an atom, yet whatever revolutions may occur in fundamental physics, the ordering of chemical atoms will remain. (Post, 1971, p. 237)

Post generalizes this example and maintains that the low-level structure of theories is particularly stable, while higher and less-confirmed levels are subject to change in the process of scientific theorizing. The pattern of the atoms remains, although quantum mechanics replaced the former framework theory. This principle seems, at first sight, to be quite plausible; but is it correct? Doubts arise once one recalls that Post himself confesses that the successful part of *S* may be smaller from the perspective of the new theory *L* than from the perspective of *S* (Post, 1971, p. 232). Given this, it is not clear how there can be a "resistant kernel" in the long run which "remains pragmatically true … for all time", as da Costa and French (1993, p. 146) suggest.

[4] For an excellent account of this case study, see Scerri (2006).

Later Post refines his proposal to also account for theories S and L with a different vocabulary. These vocabularies have to be translated into each other and this translation $\mathcal{T}$ may turn out to be more difficult than a mere one-to-one mapping. Also, a condition Q on L has to be specified such that the truncated L and S have (perhaps only approximately) the same domain. If the well-confirmed part of S is denoted by S^* (the extent of which is only a conjecture at a given time[5]) the General Correspondence Principle can be conveniently expressed as $S^*=T\,(L|Q)$ – the well-confirmed part of S is identical to the suitably translated part of L which fulfils the condition Q. If L^* is the well-confirmed part of L and S^{**} is the intersection of S^* and L^* then the thesis of zero Kuhn losses is that S^* is identical to S^{**}. Post claims that the historical record supports this thesis.[6] It should be noted, however, that Post's analysis does not take the "loser's perspective" into account. From this perspective there are indeed successes of the old theory which the new theory cannot account for.[7] Besides, even from the "winner's perspective" the thesis of zero Kuhn losses may be too strong. Saunders (1993, p. 296), for example, writes that "Laudan [(1981)] is right to insist that one can always find some theorem, deduction, conjecture, or explanation that has no precise correlate in the successor theory". He then goes on, though, to distinguish between significant and insignificant Kuhn losses; only the insignificant ones are, of course, "allowed". I will come back to this issue below. Radder (1991) has pointed out another problem for Post's approach: Not *all* equations of L may "degenerate" in equations of S. As an example, consider the famous formula $E = m_0c^2$ for the energy of a particle with rest mass m_0. This equation makes sense only in the special theory of relativity. It remains unaltered in the limit of low velocities υ (i.e., for $\beta := \upsilon/c \to 0$), although it does not correspond to an equation of classical mechanics. According to Post, the General Correspondence Principle is both a descriptive and a normative thesis. It is considered to be a *post hoc* elimination criterion and theories which do not fulfill it should be, as Post boldly advises, consigned to the "wastepaper basket" (Post, 1971, p. 235). Examining cases from the history of science, Post only spotted one "counterexample" to the General Correspondence Principle. Ironically it is the best theory we have today: quantum mechanics, a theory that, or so Post argues, does not account for the successes of its predecessor classical mechanics (Post, 1971, p. 233). This is a crucial failure which Post blames on the supposed incompleteness of quantum mechanics (Post, 1971, pp. 234, 246).[8] Quantum mechanics therefore does not, for Post, count as a case against the General Correspondence Principle. Instead the fact that quantum mechanics does not fulfil the General Correspondence Principle shows that this theory should not be accepted or at least that it should not be considered to be the successor of classical mechanics. It belongs, perhaps, in the wastepaper basket. Other proponents of a generalized correspondence principle, such as Radder, do not go as far and

5 Cf. Koertge (1973, 172 ff.).

6 For a comparison of Post's General Correspondence Principle with other correspondence principles, such as the ones suggested by Fadner, Krajewski, Radder, and Zahar see Radder (1991).

7 Cf. Hoyningen-Huene (1993, pp. 260–262) and the references to the work of Kuhn cited therein.

8 It is interesting to speculate how Post would evaluate the recent work on decoherence and the alleged "emergence of a classical world in quantum theory". See Joos et al. (2003).

emphasize correspondence relations that do hold between quantum mechanics and classical mechanics. Their arguments will be examined in the next section.

Before doing so, another issue needs to be mentioned. So far, the following three theses are in conflict: (1) Post's General Correspondence Principle is descriptively correct, (2) the belief in the truth of quantum mechanics is justified, and (3) quantum mechanics and classical mechanics share a common set of phenomena. Rather than rejecting theses (1) or (2) one might doubt thesis (3). Cartwright (1999), for example, argues that we have good reasons to believe that there are two disjunct classes of phenomena; some can be modeled by using the toolbox of quantum mechanics, others by relying on classical mechanics. There is consequently no quantum mechanical model of classical phenomena. Contrary to Cartwright, however, Post and – I believe – most physicists hold the view that quantum mechanics and classical mechanics do share a common set of phenomena. They assume that quantum mechanics accounts for the phenomena of classical mechanics *in principle*; it is merely a matter of computational complexity to demonstrate this. In the end, however, this might be nothing but a metaphysical dream.

What is the outcome of the discussion so far? First of all, when the General Correspondence Principle is applied, it often does not hold strictly, as Radder's example shows. Besides, there are losses from the loser's perspective and maybe also losses from the winner's perspective. Secondly, as a consequence of all this, there is a tension between the practice of actual science and a normative reading of the General Correspondence Principle. And yet Post is right when he points out that there is a lot of continuity in scientific theorizing, even across scientific revolutions. Still, the relations between various theories in the history of science are much more complicated than the General Correspondence Principle makes us believe. Perhaps there is no single and non-trivial principle which captures the rich structure and variety of developing scientific theories. This can only be established empirically. What is needed, therefore, is a careful examination of episodes from contemporary science and the history of science on which, perhaps, a meta-induction can be based. As a first step, it is helpful to highlight various relations which hold between successive scientific theories. This is what we will do in the next section.

3. A PLURALITY OF CORRESPONDENCE RELATIONS

In the development of scientific theories, continuities as well as discontinuities appear. Hence, the interesting question to be addressed is this: Which elements of S and L correspond to each other, and which elements do not? Are there general rules that guide practising scientists in those difficult decision situations (if it can be reconstructed as such)? As a prolegomenon to such a task, it is reasonable to examine more closely how specific scientific theories relate to each other. Which elements are taken over, what are the motives for doing so and how are the elements of the old theory made to fit the new theory? Examining cases from various sciences, I will address these questions and provide a preliminary (and not necessarily exhaustive) list of correspondence relations which *may* hold between successive theories. Some theories exhibit more than one of these relations, and some correspondences appear at different stages of the development of a theory. A first useful distinction is between *ontological* and *epistemological correspondence relations*. An ontological correspondence relation holds between *S* and *L* if some or all of the entities of *S* are also entities of *L*. In this contribution, I will consider only epistemological correspondence relations, i.e., relations between the theories in question. The following types of epistemological correspondence relations can be distinguished:

1. Term Correspondence. Here certain terms from S are taken over into L. This is a standard strategy in the development of scientific theories. In The Structure of Scientific Revolutions Kuhn writes that "[s]ince new paradigms are born from old ones, they ordinarily incorporate much of the vocabulary and apparatus, both conceptual and manipulative, that the traditional paradigm had previously employed" (Kuhn, 1996, p. 149). Now it is well-known that Kuhn also argues in the very same book that this continuity goes along with meaning variance and problems of reference. A standard example is the meaning shift from "mass" in classical mechanics to "mass" in the special theory of relativity. A disclaimer or two is in order here. Term correspondence does not imply that all terms of a theory correspond to terms in the successor theory. Often, only a few key terms are carried over, while others are left aside and new terms are coined in addition. Also, a correspondence relation between two theories can be established by a suitable translation of the respective terms. Term Correspondence is a rather minimal requirement; it is presupposed by all other correspondence relations to be discussed below.
2. Numerical Correspondence. Here *S* and *L* agree on the numerical values of some quantities (cf. Radder, 1991, pp. 203–204). Numerical Correspondence therefore presupposes Term Correspondence. An example is the spectrum of hydrogen in the Bohr model and in quantum mechanics. Although the assumptions that were made to calculate the spectrum differ considerably in both theories, they nevertheless lead to the same numerical values. Again, this is a rather weak kind of a correspondence relation which is moreover usually realized only approximately (as in the example just discussed). Its heuristic value is low since the principle can only be applied *post hoc*. Obviously, Numerical Correspondence is only interesting in the mathematical sciences; it does not apply, for instance, in large parts of biology or archeology.
3. Observational Correspondence. This kind of correspondence relation is introduced in Fine (1993) in the context of his interesting resolution of the quantum mechanical measurement problem. Fine does not accept Cushing's claim that Bohm's version of quantum mechanics should have been chosen according to Post's General Correspondence Principle (Cushing, 1993, p. 262), because the Bohm theory "did not enable one to retrieve the classical and well-confirmed account of a ball rebounding elastically between two walls" (Fine, 1993, p. 280). It therefore does not fulfil Post's correspondence principle. Bohm's theory does, however, fulfil a weaker form of a correspondence principle. Fine writes: "[W]here the classical account itself is well-confirmed, the Bohm theory 'degenerates' into the classical account of what we are expected to observe under well-defined conditions of observation"

(Fine, 1993, p. 280). Unfortunately, the standard Copenhagen version of quantum mechanics does not fulfil the principle of Observational Correspondence and Fine therefore presents his solution of the measurement problem in order to restore this. Abstracting from quantum mechanics, Observational Correspondence means that L "degenerates" into what we are expected to observe according to S* under well-defined conditions of observation. Observational Correspondence, like Numerical Correspondence, presupposes Term Correspondence, but differs from Numerical Correspondence, which may also apply when the quantities in question cannot be observed. Besides, Observational Correspondence relations can also hold in sciences which do not represent their content numerically. Observational Correspondence emphasizes the role of the conditions of observation which are especially important in the context of quantum mechanics. A heuristic principle based on the demand of Observational Correspondence is again only a post hoc selection criterion. It is of no help in the actual process of constructing new theories. Observational Correspondence alone also does not suffice to provide an explanation for the success of the old theory. It is therefore weaker than Post's General Correspondence Principle.

4. Initial or Boundary Condition Correspondence. According to a well-known view of scientific theories, a theory is a set of axioms (or laws) plus suitable initial or boundary conditions. Kamminga (1993) complains that the philosophical focus is too much on the axioms (or laws), leaving initial and boundary conditions aside. This is unfortunate, since especially in the non-formal sciences, Kamminga claims, these conditions play an important role which is relevant to the issue of inter-theory relations. It turns out that there are theories which incorporate consequences of their predecessor as an initial or boundary condition. An example from the research on the origin of life illustrates Kamminga's general point which she sums up as follows: "[I]n the attempt to integrate the original theory *T* with another theory outside its domain, some consequence of the latter is incorporated into *T* as an antecedent condition, which then places strong constraints on the selection of laws that have explanatory relevance in the modified theory *T*'" (Kamminga, 1993, p. 77). This procedure, therefore, provides a link between the two theories. Note, however, that this way of connecting two theories is only a very loose one. It has some heuristic value but it should be noted that the assumptions taken over from the predecessor theory remain unexplained in the successor theory.
5. Law Correspondence. Laws from *S* also appear in *L*. This kind of correspondence relation often holds only approximately. An example are the laws for the kinetic energy in classical mechanics and in the special theory of relativity. For low velocities, $T_{CM} = 1/2\ m\upsilon^2$ and $T_{SRT} = (m - m_0)\ c^2 = 1/2\ m\upsilon^2 \cdot (1 + 3/4\ \beta^2 + O\ (\beta^4))$ are approximately the same. Hence, the special theory of relativity reproduces and explains the successful part of classical mechanics. It is probably this kind of a correspondence relation which Post had in mind when he suggested his General Correspondence Principle. Law Correspondence implies Numerical Correspondence and presupposes Term Correspondence, the difficulties of which (such as meaning variance, etc.) therefore occur again. Despite all this it is required that the terms in question have the same operational meaning in *S* and *L* (cf. Fadner, 1985,

p. 832). In many cases, Law Correspondence is only a *post hoc* selection criterion of theory choice. As Radder's above-mentioned example demonstrates, it may only hold for *some* of the laws of the theories in question.

6. Model Correspondence. This type of a correspondence relation comes in two variants. (1) A model which belongs to *S* survives theory change and reoccurs in *L*. A typical example is the harmonic oscillator which is widely used in classical mechanics, but is also applied in quantum mechanics and in quantum field theory. It should be noted that models, such as the harmonic oscillator, are not only taken over by the theory which succeeds the original theory, but also by quite unrelated theories. This is best seen by pointing to all other theories of physics which employ the harmonic oscillator; in fact, it is difficult to find a theory which does not employ this model. Model Correspondence of this first kind has a considerable heuristic potential. It is, however, not guaranteed that the new theory explains the success of the old theory, because the model in question may be embedded in a completely new framework theory which may also affect the overall correspondence relation between *S* and *L*. (2) Post mentions another strategy of theory construction which takes models seriously: "In this case we adopt a model already available which may initially have been offered as an arbitrary articulation of the formalism only. […] It is a case of borrowing a model of the *S*-theory which contained features not *essential* for the modelling of the *S*-theory ('neutral analogy'), and assigning physical significance to such extra features" (Post, 1971, p. 241). An example is the crystallographic models which were used already a century before physicists identified the units of the regular lattices with physical atoms. Sometimes, Post concludes, scientists built "better than they knew" (Post, 1971, p. 242). This example also shows that Model Correspondence of this second kind may indeed lead to an explanation of the success of the predecessor theory.[9] However, the criterion is highly fallible, as Post himself grants.
7. Structure Correspondence. Here the structures of *S* and *L* correspond. But what is a structure, and what does it mean that two structures correspond? One option is to use the term "structure" only in its precise mathematical meaning. If one does so, it is not difficult to flesh out the idea of a correspondence relation between two structures by applying mathematical concepts such as sub-groups and group contractions. And indeed, many theories in physics can be linked to each other in this way. A typical example is the relation between the inhomogeneous Lorentz group (that characterizes the special theory of relativity) and the inhomogeneous Galilei group (that characterizes classical mechanics) which "correspond" in a precise mathematical sense. In examples like this Structure Correspondence works best. Another interesting case is the relation between the theories of Ptolemy and Copernicus. Saunders shows that "[a]n astronomy based only on epicycles […] corresponds to an expansion of the form $\Sigma_i\, c_i \exp(i\omega_i t)$ (with the earth chosen as origin)" (Saunders, 1993, p. 299). So the mathematical structures of both theories are (perhaps only approximately) the same, which leads Saunders to the conclusion

[9] For more on the relation between models and theories see Frigg and Hartmann (2006).

> that there is no reason to worry about the abandonment of the Aristotelian world-view or a wholesale change of paradigm (Saunders, 1993, p. 298).

Saunders' large-scale fight against relativism[10] appears somewhat cheap; parts of theories where problems for the idea of correspondence show up are stamped as "insignificant" (such as the ontology of a theory,[11] but also laws) while only the mathematical structure of a theory remains, in some sense, stable. But even here things are not that easy. With respect to the role of gravity, Saunders concedes that he does "not suggest that these things can be completely codified" but goes on to confess that this strategy "is, and [. . .] has always been, the essence of the enterprise of dynamics" (Saunders, 1993, p. 306). Confessions like this are not enough to provide a vigorous defence of the cumulative, progressive view of the history of physics, but Saunders showed that mathematical structures of consecutive theories may and often do correspond in a strict mathematical sense.

Structure Correspondence does not imply Numerical Correspondence. Often, the structure is "too far away" from the empirical basis of a theory in order to guarantee continuity at that level (especially in the cases Saunders has in mind). It is therefore not at all trivial to reproduce the empirical success of the precursor theory once one has decided to take over parts of the structure of the old theory. Despite this, Structure Correspondence has a very high heuristic value, especially in contemporary theoretical physics. Because of the huge gap between the respective theories (such as superstring theory) and the world to which we have empirical access, structure-based reasoning, such as symmetry considerations, is often the only tool which enables scientists to proceed. However, when it comes to other sciences, such as biology or archeology, Structure Correspondence does not seem to be of much value, especially if one explicates 'structure' mathematically.

Three conclusions can be drawn from the above analysis: First, successive theories can be related in many ways. Sometimes only Numerical Correspondence holds (at least approximately), in other cases entire mathematical structures correspond. Every case is different. This is why global philosophical issues, such as meaning variance and incommensurability, should be discussed "locally", i.e., on the basis of concrete case studies that exemplify specific types of relations between scientific theories. The details might, and often do, matter.

Second, there are continuities and discontinuities in scientific theorizing, although it is not a priori clear which elements of a theory will survive theory change, and which ones will have to go. An additional difficulty for correspondence theorists is the notorious problem of underdetermination. Maybe there is no unique best choice regarding which elements of successive theories should correspond and which should not correspond with each other.

Third, the philosophical project of a methodology is best described by the picture of a *toolbox*. According to this view, methodologists extract – ideally on the basis of a wealth of case studies – a set of methods and techniques which can tentatively be applied by practicing scientists in a particular situation. What is in the toolbox may,

[10] Note that for Saunders, "relativism" is a collective name for social constructivism, historicist epistemology, linguistic holism and anti-realism; cf. Saunders (1993, 295 ff.).

[11] Cf. Saunders' discussion of ether in Saunders (1993, p. 299).

however, depend on time. Good scientists know, of course, already a wealth of tricks and methods, and they also know how to use them flexibly and appropriately. This view of the status of methodology is a middle ground between two extreme positions. Zahar (1983, 258 ff.) defends a rather strong form of a rational heuristics which leaves little room to chance and other influences, while Popper's (1972, Chap. 7) evolutionary picture supports the opposite view, that there is no rational heuristics and it is the job of the scientists to make bold conjectures which then have to "survive" empirical tests and rational criticism (cf. Radder, 1991, 201 ff.). My conclusion seems, after all, to be similar to Post's view on the role of heuristics which he illustrates with an apt analogy: "The study of the structure of existing houses may help us in constructing new houses" (Post, 1971, p. 217).

4. MODELLING HIGH-TEMPERATURE SUPERCONDUCTIVITY: A CASE STUDY

In this section, I will look at a case study from contemporary physics and ask which role heuristic principles such as the General Correspondence Principle play in science. The case study deals with a scientific episode that is ongoing. So far, there is no consensus in the scientific community, only a multitude of more or less elaborated competing research programs and strategies. I'll identify some of these strategies and ask which role correspondence considerations play when scientists are confronted with an intricate problem.

The case study deals with our theoretical understanding of high-temperature superconductivity. Conventional superconductivity is a phenomenon long well known and understood. It occurs at extremely low temperatures close to the absolute zero. For a long time, it was considered to be impossible to find or produce substances that remain superconducting if the temperature is raised to, say, 30 K or more. The breakthrough occurred in 1986 with the work of Georg Bednorz and Alex Müller, who discovered the first high-temperature superconductor LBCO with a transition temperature of 35 K (Bednorz and Müller, 1988). Later, materials with a transition temperature of up to 160 K were produced. It turned out that the materials that are superconducting at such high temperatures are very special: they have a layered structure made up of one or more copper–oxygen planes and exhibit an abnormal behavior also when in the normal state. This complicates the theoretical understanding of these so-called cuprates considerably and so it is no surprise that, despite a lot of theoretical work over the last 20 years and a wealth of experimental data, no theoretical understanding of high-temperature superconductivity is forthcoming.[12]

There is, however, a well-confirmed theory of conventional superconductors. This theory, the so-called BCS theory – named after its inventors James Bardeen, Leon Cooper and Robert Schrieffer – is a microscopic theory that explains the appearance

[12] See Tinkham (1996) and Waldram (1996) for (somewhat dated) overviews. More recent reviews are Anderson (2006) and Kivelson (2006).

of a superconducting phase by showing how bound pairs of electrons with opposite spin and momentum are formed if the temperature is below the critical temperature. Although the (fermionic) electrons in a so-called Cooper pair are often separated from each other by a large distance, they act effectively as a single particle which has the quantum statistical properties of a boson. And this is why a large number of Cooper pairs can be in the lowest energy state, which in turns leads to the vanishing of the electrical resistance in the materials. Complementing the BCS theory, Bardeen, Cooper and Schrieffer proposed a specific model – the BCS model – that specifies the concrete mechanism that is responsible for the creation of Cooper pairs. This mechanism is based on so-called *s*-wave interactions of the electrons, mediated by the vibrating crystal lattice, and accounts for all phenomena involving superconductivity discovered until 1986.[13]

When it comes to understanding high-temperature superconductivity, two points are uncontroversial: (i) The BCS theory also applies, i.e., high-temperature superconductivity results from the formation of electron pairs at temperatures below the (material-dependent) critical temperature. (ii) The BCS model does not apply. More specifically, it is generally accepted that *s*-wave interactions cannot account for the extremely high transition temperature that we find in the cuprates. And so the task is to develop a new model. To do so, physicists follow a wide variety of approaches that can be located on a spectrum ranging from conservative to highly revolutionary apoproaches.

Conservative approaches aim at developing an account that deviates as little as possible from the theoretical framework of the BCS theory (i.e., the Fermi liquid theory) and the more specific assumptions of the BCS model. *Revolutionary* approaches attempt to formulate an account of high-temperature superconductivity in a new theoretical framework and propose mechanisms that deviate considerably from the BCS model. While some authors suggest different mechanisms for different types of materials, others want to identify *the* mechanism of high-temperature superconductivity. Besides these extreme approaches, a whole range of approaches in between has been put forward.

All of these approaches are developing, constantly modified and occasionally completely rejected, but none has yet succeeded. Even 20 years after the astonishing discovery of Bednorz and Müller, there is no satisfactory and generally accepted theory of high-temperature superconductivity. Given that we had to wait 46 years from the discovery of conventional superconductivity (by Heike Kamerlingh Onnes in 1911) to the formulation of the BCS theory (in 1956), we might have to be patient for quite a while. That it is taking so long is seen as an indication that a major theoretical breakthrough can be expected. In situations like this, much is at stake and the debate among the members of the relevant scientific community often touches philosophical and methodological issues. So let us have a look at some of the proposals that are currently discussed:

[13] For technical details, see Tinkham (1996, Chap. 3).

4.1. The conservative strategy

The proponents of a conservative research strategy (such as David Pines and David Scalapino) propose rather minimal modifications of the BCS model in order to account for the phenomenon of high-temperature superconductivity. They keep, for example, the Fermi liquid theory as a theoretical framework. What is replaced, in a minimal way, is the specific pairing mechanism. Instead of *s*-wave interactions among the electrons, the next more complicated interactions, mediated by *d*-waves (plus spin fluctuations), are considered. *d*-waves are suggested by the spatial structure of the substances in question and so have an empirical foundation (Scalapino, 1995). Supporters of this approach (and related approaches) point to its successes when it comes to understanding experiments. They also stress the minimal character of the modifications on the BCS model that have to be made.

Critics (such as Philip Anderson) object that the spin-fluctuation theory is not a real theory, but just a bunch of heuristic tools with various unjustified assumptions. It also counts against the spin-fluctuation theory, or so Anderson claims, that it can only be applied by using high-powered computer simulations – a procedure which cannot be controlled independently. Supporters reply that the complexity of the problem suggests that it is treated with such methods.[14]

More recently, a peculiar charge ordering pattern in two separate states and regimes of the cuprates has been discovered. To explain these patterns, S. C. Zhang used the crystalline ordering of the *d*-wave Cooper pairs and showed how this new state of matter fits into the global phase diagram of cuprates with a high transition temperature. He also derived several predictions from his theory, which appear to have recently been verified in transport measurements. Indeed, in the last decade the experiments showed more clearly that the cuprates with a high transition temperature have a very complex phase diagram with many competing phases. In optimally and underdoped materials, the state of matter out of which the superconductivity arises exhibits a so-called *pseudogap* at temperatures which are high compared to the transition temperature (Norman et al., 2005). The pseudogap was found only 3 years after Bednorz and Müller's original discovery, but its physical origin, its behavior and whether it constitutes a distinct phase of matter is still not well understood. What is clear, however, is that the pseudogap will play an important role in any theory of high-temperature superconductivity. Lee (2006), for example, believes that the existence of the pseudogap supports the Charge Density Wave (CDW) theory as the large pseudogap in the cuprates can be generated by a CDW order with *d*-symmetry.

4.2. The revolutionary strategy

Philip Anderson started off defending a truely revolutionary strategy. He submited that the problem of understanding high-temperature superconductivity cannot be attacked by a minimal modification of the BCS model. He hence proposed to give up a standard

[14] See Anderson (1995), Anderson and Schrieffer (1991), Anderson and Mott (1996) and Ford and Saunders (1997).

assumption in solid state physics – that the system can be described as a Fermi liquid – and to replace it with the assumption that the systems in question are, in the normal as well as in the superconducting state, so-called Luttinger liquids. Anderson's account was highly speculative and was rejected by most physicists.[15]

Meanwhile, in the light of new experimental evidence, Anderson gave up his original account as well and proposed a new one with a somewhat different flavor (Anderson et al., 2004). This new account is based on an interesting analogy between a Mott insulator and high-temperature superconductivity and is sometimes called a theory of a doped Mott insulator. Though less revolutionary than the old one, Anderson's new account is fundamentally different from other approaches that are mostly based on perturbative many-body theory. During the last 2 years, it has attracted considerable attention amongst both theorists and experimentalists. However, it is still far from being generally accepted.

And so the question remains how conservative or revolutionary a future theory of high-temperature superconductivity should be. As there is no consensus in the scientific community over which strategy is the right one, a final assessment of the various strategies is not possible at the moment. While a majority of physicists seem to be in favour of a conservative account, Anderson's (1995, p. 38) reply is still worth taking seriously: "[I]f it is time for a revolution, enjoy it and relax!"

5. CORRESPONDENCE AT BAY? – SOME LESSONS

What are the implications of our case study for the General Correspondence Principle? First, it has to be noted that the starting point for the development of theories of high-temperature supercondictivity was not an internal anomaly of the predecessor theory, i.e., the BCS theory (and the BCS model) of conventional superconducters. Quite to the contrary, the account given by BCS turned out to be an excellent tool to learn about the physics (and chemistry) of superconductors for many years. No one doubted, for example, that it is a perfectly consistent theory. And yet after a while, new experimental data were produced that could not be fit into the BCS account.[16] This is what prompted the current developments.

Second, none of the proposals for a theory of high-temperature superconductivity that have been for put forward so far contain the BCS account as a limiting case. There is, hence, no new theory that accounts for the successes of the old theory as well as for the class of new phenomena. And so the General Correspondence Principle seems to be violated. Or isn't it? A way out for the correspondence theorist is to argue that the General Correspondence Principle cannot be applied to the case of high-temperature superconductivity. She could argue that the corresponding materials belong to a completely different class of materials, that the phenomena in question (i.e., conventional and high-temperature superconductivity) are too different, and

[15] See Anderson (1997) for a book-length exposition of his approach and Leggett (1997) for a criticism.
[16] The BCS account is the BCS theory plus the BCS model.

that the similarity between them is at best superficial. And so we are not supposed to apply the General Correspondence Principle (or something like it). What is at stake is the question of what the domain of applicability of methodological principles such as General Correspondence Principle is and how we find or identify this domain. I have no principled answer to this question (and doubt that there is one). At best, I think, we can trust the judgement of the scientific community, which is the only authority I can see. For now, the scientific community has not made up its mind on high-temperature superconductivity, and the field is characterized by a large amount of dissent over questions ranging from specific scientific questions to methodological questions.

Looking more closely at the various proposals for a theory of high-temperature superconductivity, one realizes that all of them take over some elements of the BCS account. All of them, for example, adopt the idea that pairing is responsible for superconductivity.[17] And some of them modify the BCS pairing mechanism only slightly, so that the mechanism responsible for high-temperature superconductivity can be understood as a natural extension of the mechanism for conventional superconductivity. I conclude from this that there is certainly some continuity on the theoretical level. But continuity is not an all-or-nothing matter (as the General Correspondence Principle makes us think). How far the continuity reaches (or should reach) will depend on the specific case in question and is a matter of debate. In the course of science, different (and sometimes even the same) scientists will try out different approaches and adopt different general methodological or scientific principles. The best way to understand this, as already argued at the end of section 3, is to consider methodological principles as *tools* that are tentatively adopted by some scientists for a specific purpose. Whether they are good tools or bad tools depends on the specific case and will be decided, in the end, by the scientific community.

The General Correspondence Principle is just one tool among many that scientists tentatively use in the process of theory (or model) construction. In many cases, there will be a lot of dissent over the question of how closely the successor theory should be related to its predecessor. The respective considerations are typically subtle and deeply entangled with the problem at hand. I submit that general methodological principles unfold their fruitful action only when adapted to a specific problem.

This raises a difficulty for those philosophers of science who are interested in the identification of general methodological principles. Instead of doing this, it might be best to restrict oneself and help put together a *toolbox of principles* that scientists can tentatively use on an as-needed basis. To do this seems to me to be an important task. It is both doable – unlike the program of rational heuristics – and worthwhile. Philosophers of science are especially qualified to undertake such a project, as they overlook a greater part of science than do practitioners, who are often confined between the boundaries of their field of specialization.

[17] For a recent discussion, see Tsai and Kivelson (2006).

6. RATIONALITY, REALISM AND COHERENCE

Even if the continuity that we observe in scientific theorizing does not fit into the neat picture that goes with the Generalized Correspondence Principle, there is no doubt that there is a lot of continuity (or stability, as Hacking puts it) in scientific theorizing. So the following questions arise: Why is it that there is so much continuity in science? And how can this continuity be understood philosophically? Note that these are questions that cannot be attacked locally, i.e., on the basis of case studies alone.

In "The Social Construction of What?", Hacking (1999) presents two explanations for the prevelent continuity in scientific theorizing. The *first* explanation, realism, stresses factors internal to science and comes in various variants. One of them is convergent realism, which holds that successive theories of "mature science" approximate the truth (i.e., the ultimate or final theory) better and better. This presupposes the existence of a measure of the distance of a given theory from the truth (or at least an ordering relation), which is a controverial topic despite all the worthwhile work on verisimilitude and truthlikeness,[18] and conflicts with the many discontinuities that emerged in the development of "mature" scientific theories, as Laudan (1981) has convincingly demonstrated. Another problem of convergent realism is that there might be no ultimate theory. It is possible that the process of constructing ever better theories never ends because "there are infinitely many levels of structure that can be unpacked, like an infinitely descending sequence of Chinese boxes, or to use the more colloquial expression: it is boojums all the way down" (Redhead, 1993, p. 331).

A weaker variant of realism, which seems to be able to handle the problems raised by Laudan and others, is structural realism. According to (one version of) this position, at least the high-level mathematical structures of scientific theories converge. Continuity on the level of ontology and perhaps even on the level of one or another law is not required. However, my discussion in the previous section suggests that there is not enough "empirical evidence" to support this position. Besides, there are more plausible ways to explain the persistence of certain mathematical structures; I will come back to this below.[19]

The *second* explanation discussed by Hacking, constructionism, comes in several variants as well. All of them emphasize the role of factors external to science. For example, in the previously-mentioned section on stability, Hacking quotes the historian of science Norton Wise, who argues that culture and science are inseparably connected with each other. Cultural influences go into the discovery of scientific theories and leave an indelible footprint there. Steven Weinberg, whom Hacking quotes approvingly, argues, however, that these influences "have been refined away" (Hacking, 1999, p. 86).

Adopting a Kuhnian line of thought, a constructionist can explain the remarkable stability of scientific theories as follows: she reminds us that scientists grow up and get trained in a certain research tradition, they learn all the theories and techniques of

[18] See Niiniluoto (1999) for a recent exposition.

[19] Another realist way put forward to account for the stability of scientific theories is Radder's (1988, 1991) moderate realism.

this tradition and, of course, they want to promote their career. These scientists are well-advised to demonstrate their affiliation to the tradition in question by contributing to the corresponding research program. All too radical junior-scientists are not protegéed and their careers may take a turn for the worse. After all, the scientific community does not reward disloyal behaviour. Here is another, and perhaps somewhat more plausible, explanation of the continuity in scientific theorizing by external factors: It is simply too costly to start from scratch when confronted with a new scientific problem. Scientists who do so will not be able to produce a sufficient number of publications in renowned journals which is essential to carry on with academia. This also explains why the high-level structure of theories is so stable: since so much depends on it, a revision or replacement would be very costly indeed. Although there is some undeniable truth to these stories, I think that there is more to be said.

I conclude that both explanations, realism and constructionism, do not suffice to account for the prevalent continuity of scientific theorizing. Interestingly, they do not account for the occasional discontinuities (or revolutions, as Hacking puts it) in scientific theorizing as well. Clearly, discontinuities do not fit in the scheme of convergent realism. But also weaker forms of realism do not explain the *relative weights* of continuities and discontinuities in scientific theorizing. And constructionism fails to do so as well. However, it is the (possibly case-dependent) relative weights that cry out for an explanation and so I conclude that an alternative explanation is needed that helps us understand why there is as much continuities and as much discontinuities in scientific theorizing as there is. Such an explanation should fulfill the following three conditions. (i) It should be sufficiently general, as our goal is a philosophical understanding of the scientific practice. (ii) It should be descriptively correct and account for successful scientific practice. (iii) It should combine internal and external factors, as both seem to be important.

Providing an explanation of the relative weights of continuities and discontinuities in scientific theorizing that fulfills these three conditions is a difficult task and an easy answer such as "realism" or "constructionism" is unlikely to succeed. Instead I suggest to adopt a formal philosophical framework and construct *philosophical models* within this framework, mimicking the successful scientific methodology of constructing models of a theory mentioned in the Introduction. The philosophical framework will be general and satisfies condition (i). The models will provide details to account for more specific cases, which helps satisfying condition (ii). Constructing more specific models makes sense as the relative weights will depend on the specific scientific theories in question. The framework we chose will make sure that internal and external factors are taken into account, thus satisfying condition (iii). Let us now see how this works.

The central idea of my own explanation, coherentism, is that the transfer of elements of the predecessor theory S (whose well-confirmed part S^* is non-empty) into the successor theory L increases the overall coherence of our knowledge system. The word "coherence" is here used in the epistemological sense, i.e. as a measure for how well a body of beliefs "hangs together". As Lawrence BonJour explains, "[it] is reasonably clear that this 'hanging together' depends on the various sorts of inferential, evidential, and explanatory relations which obtain among the various members of a system of belief, and especially on the more holistic and systematic of these." (BonJour, 1985, p. 93) Given this account of coherence, it is plausible that S (or better: S^*) and L cohere more if, ceteris paribus, the two theories share common elements.

Before working out my coherentist account, a potential objection has to be addressed. One may ask why both S (or S^*) and L should be in our knowledge system at the same time. Why isn't S simply abandoned and replaced by a theory L which is itself more coherent than the conjunction of S^* and L? Although some theories in the history of science were indeed abandoned altogether and no elements were transfered from S to L(such as the phlogiston theory), this is often (and especially with the theories of the last 150 years)[20] not the case. Theories such as classical mechanics are still successfully applied, although this theory is, strictly speaking, superseded by quantum mechanics and Einstein's theories of relativity. There are indeed a host of reasons why we still work with (the well-confirmed part of) classical mechanics. Most importantly, perhaps, is that classical mechanics has the practical advantage that it can be applied to problems where, for instance, quantum mechanical calculations are not feasible (and would presumably give the same result anyway). There is, however, also an epistemic reason why classical mechanics is still so popular. The theory provides explanations that quantum mechanics is unable to provide. This is, again for practical reasons, because classical mechanics can account for phenomena quantum mechanics cannot account for. However, we also favor the Newtonian explanation of, say, the tides over an explanation based on Einstein's general theory of relativity as the concepts classical mechanics employs (such as force) are "closer" to the phenomenon in question which arguably helps us to understand. An explanation of the tides based on the properties of curved spacetime would only obscure the situation. Good explanations function like this, and as it is one of the goals of the scientific enterprise to provide explanations of the phenomena in its domain, (the well-confirmed part of) classical mechanics is indispensable and remains an integral part of our knowledge system.

Let us now develop our tentative account of scientific theory change. To do so, we need (i) a precise account of the notion of coherence and how it is measured. To measure the coherence of a scientific theory (or of a set of scientific theories), we also need (ii) a representation of a scientific theory that suits this request. As we will see, choosing a Bayesian framework will satisfy both requests.

We explicate (the empirical part of) a scientific theory T as a set of interrelated models $\{M_1,\ldots,M_n\}$. Here a model is represented by a proposition M_i which may itself be a conjunction of more elementary propositions (such as instantiations of laws, additonal model assumptions and initial and boundary conditions). Each model M_i is related to a phenomenon, represented by a proposition E_i (for "evidence"), that it accounts for. We define a probability distribution P over all model variables M_i and all phenomena variables E_i. This probability distribution is best represented by a Bayesian Network. In such a network, there is an arrow from each model node M_i to its corresponding phenomenon node E_i, and there are various arrows between the model nodes which reflects the idea that the models in a scientific theory mutually

[20] Cf. Fahrbach (forthcoming).

support each other. By doing so, BonJour's "various sorts of inferential, evidential, and explanatory relations which obtain among the various members of a system of belief" can be modeled probabilistically. For example, if model M_2 entails model M_1, then $P(M_1|M_2) = 1$, and if model M_2 supports model M_1, then $P(M_1|M_2) > P(M_1)$ (for details, see Hartmann, 2007).

In the next step, these probabilistic dependencies between models have to be aggregated to arrive at a coherence measure $coh(T)$ of the theory T. While there are several different proposals in the literature for how to do this, my preferred one is laid out in my book *Bayesian Epistemology* (with Luc Bovens). In this book, we show that no such coherence measure exists and that all that can be done is the specification of a function that generates a partial ordering over a set $\mathcal{T} = \{T_1,\ldots,T_n\}$ of theories. Sometimes there is no fact of the matter, which of two theories is more coherent. But often there is.

At this point, the natural question arises why coherence is a good thing. Even if scientists do aim for a highly coherent knowledge system (which explains the use of correspondence considerations), it is not clear why science *should* aim for a highly coherent knowledge system. So what is at stake is the normative question. Simply put, the answer is that, given certain conditions, coherence is truth-conducive.[21] In the Bayesian framweork, this means that the more coherent set of, say, models, is also the one that has the greater posterior probability (provided that certain ceteris paribus conditions are satisfied). This is an important result as it makes informal discussions of the coherence theory of justification more precise (Bovens and Hartmann, 2003a, b).

Similarly, the following can be proven: If the new theory L adds one additional model to its predecessor S^* and if L is more coherent than S^*, then L has a higher posterior probability than S^* if the prior probability of L is not much lower than the prior probability of S^*. For details and a proof, see Hartmann (in preparation). Note that there is clearly a correspondence relation between S^* and L as all successes (as well as the failures) of the old theory are taken over in the new, more encompassing theory.

What I have presented are the initial steps of a philosophical research program. Clearly much more needs to be said and more philosophical models have to be studied to better evaluate the prospects of a full-fledged Bayesian account of scientific theory change. For example, a Bayesian account of the various correspondence relations discussed in Sect. 3 has to be given. If all goes well, we have a methodology at hand that will help us to explain (using the notion of coherence) and justify (using the truth-conduciveness of coherence) the successful scientific practice. But we have to be cautious and not expect too much. As Earman (1992, Chap. 8) points out, scientific theory change presents a whole range of other problems for the Bayesian. So I do not pretend that *all* aspects of scientific theory change can be accounted for in Bayesian term. However, until the contrary is proven, I follow Salmon (1990) and take Bayesianism to be an attractive philosophical research program that helps us to illuminate the intricacies of scientific theory change.

[21] I here follow the standard usuage of the term. Note, however, that "truth-conducive" is a misnomer as Bayesians only specify subjective probabilities, and a posterior probability of 1 does not entail the truth of the proposition in question. I thank Kevin Kelly for pointing this out to me.

Before closing, a word is in order about the nature of the probabilities in our Bayesian framework. Bayesians typically assume that individual agents have subjective degrees of belief which I take to include internal as well as external factors. These degrees of belief may differ from scientist to scientist as different scientists may have different background knowledge and different "scientific tastes". And yet, we observe that different scientists often assess new hypotheses similarly and reach similar conclusions about their acceptability. There is much to Kuhn's idea that the scientific community is the real decision-making agent, and it is an interesting research project to construct models for the aggregation of individual probability functions to a probability function of the scientific community. I have to leave this task for another occasion.

Acknowledgements

I am indebted to Léna Soler and Jan Sprenger for excellent comments and suggestions, to Clark Glymour for very useful discussions and feedback, and to Maria Elena Di Bucchianico for help with the case study.

BIBLIOGRAPHY

Anderson, P. and Schrieffer, R. (1991) A Dialog on the Theory of High-T_c. *Physics Today*, June, 55–61.

Anderson, P. (1995) Condensed Matter: The Continuous Revolution. *Physics World*, 8, December, 37–40.

Anderson, P. and Mott, N. (1996) High-Temperature Superconductivity Debate Heats Up. *Physics World*, 9, January, 16.

Anderson, P. (1997) *The Theory of Superconductivity in the High-T_c Cuprates*. Princeton NJ: Princeton University Press.

Anderson, P., Lee, P., Randeria, M., Rice, T., Trivedi N., and Zhang, F. (2004) The Physics Behind High-Temperature Superconducting Cuprates: The 'Plain Vanilla' Version of RVB. *Journal of Physics – Condensed Matter*, 16(24), R755–R769.

Anderson, P. (2006) Present Status of the Theory of the High-T_c Cuprates. *Low Temperature Physics*, 32 (4–5), 282–289.

Bednorz, J. G. and Müller, K. A. (1988) Perovskite-Type Oxides – The New Approach to High-T_c Superconductivity. *Reviews of Modern Physics*, 60, 585–600.

BonJour, L. (1985) *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.

Bovens, L. and Hartmann, S. (2003a) *Bayesian Epistemology*. Oxford: Oxford University Press.

Bovens, L. and Hartmann, S. (2003b) Solving the Riddle of Coherence. *Mind*, 112, 601–634.

Cartwright, N. (1999) *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Cushing, J. (1993) Underdetermination, Conventionalism and Realism: The Copenhagen vs. the Bohm Interpretation of Quantum Mechanics. In S. French and H. Kamminga (1993), pp. 261–278.

da Costa, N. and French, S. (1993) Towards an Acceptable Theory of Acceptance: Partial Structures and the General Correspondence Principle. In S. French and H. Kamminga (1993), pp. 137–158.

Earman, J. (1992) *Bayes or Bust? An Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT.

Fadner, W. (1985) Theoretical Support for the Generalized Correspondence Principle. *American Journal of Physics*, 53(9), 829–838.

Fahrbach, L. (forthcoming) The Pessimistic Meta-Induction and the Exponential Growth of Science. Unpublished manuscript. University of Konstanz.

Fine, A. (1993) Measurement and Quantum Silence. In S. French and H. Kamminga (1993), pp. 279–294.

Ford, P. and Saunders, G. (1997) High Temperature Superconductivity – Ten Years On. *Contemporary Physics*, 38, 63–81.

French, S. and Kamminga, H. (eds.) (1993) *Correspondence, Invariance and Heuristics. Essays in Honour of Heinz Post*. Dordrecht, The Netherlands: Kluwer.

Frigg, R. and Hartmann S. (2006) Models in Science. *Stanford Encyclopedia of Philosophy* (Spring 2006 Edition).

Hacking, I. (1999) *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Hartmann, S. (2002) On Correspondence. *Studies in History and Philosophy of Modern Physics*, 33B, 79–94.

Hartmann, S. (2008) Modeling in Philosophy of Science, to appear in: M. Frauchiger and W.K. Essler (eds.), *Representation, Evidence, and Justification: Themes from Suppes (Lauener Library of Analytical Philosophy; vol.* 1). Frankfurt: ontos verlag.

Hartmann, S. (in preparation) Normal Science and Its Justification.

Hoyningen-Huene, P. (1993) *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. Chicago IL: University of Chicago Press.

Joos, E. et al. (2003) *Decoherence and the Appearance of a Classical World in Quantum Theory*. Berlin: Springer.

Kamminga, H. (1993) Taking Antecedent Conditions Seriously: A lesson in Heuristics from Biology. In S. French and H. Kamminga (1993), pp. 65–82.

Kivelson, S.A. (2006) Superconductivity on the Verge of Catastrophe. *Nature Materials*, 5(5), 343–344.

Koertge, N. (1973) Theory Change in Science. In G. Pearce and P. Maynard (eds.) *Conceptual Change*. Dordrecht, The Netherlands: Reidel, pp. 167–198.

Kuhn, T. S. (1996) *The Structure of Scientific Revolutions*. Chicago IL: University of Chicago Press.

Laudan, L. (1981) A Confutation of Convergent Realism. *Philosophy of Science*, 48, 19–49.

Lee, D.H. (2006) D-Symmetry CDW, Fermi Arcs, and Pseudogap in High T_c Cuprates, talk presented at *The 8th International Conference on Materials and Mechanisms of Superconductivity and High Temperature Superconductors*. Dresden, 9–15 July 2006.

Leggett, T. (1997) Superconducting Thoughts Meet Sceptical Resistance. *Physics World*, 10, October, 51–52.

Niiniluoto, I. (1999) *Critical Scientific Realism*. Oxford: Oxford University Press.

Norman, M., Pines, D. and Kallin, C. (2005) The Pseudogap: Friend or Foe of High T_c? *Advances in Physics*, 54(8), 715–733.

Popper, K. (1972) *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon.

Post, H. (1971) Correspondence, Invariance and Heuristics. *Studies in History and Philosphy of Science*, 2, 213–255. In S. French and H. Kamminga (1993), 1–44.

Radder, H. (1988) *The Material Realization of Science*. Assen: Van Gorcum.

Radder, H. (1991) Heuristics and the Generalized Correspondence Principle. *British Journal for the Philosophy of Science*, 42, 195–226.

Redhead, M. (1993) Is the End of Physics in Sight? In S. French and H. Kamminga (1993), pp. 327–341.

Salmon, W.C. (1990) Rationality and Objectivity in Science, or Tom Kuhn Meets Tom Bayes. In C. W. Savage (ed.) *Scientific Theories*. Minneapolis MN: University of Minnesota Press, pp. 175–204.

Saunders, S. (1993) To What Physics Corresponds. In S. French and H. Kamminga (1993), pp. 295–325.

Scalapino, D. (1995) The Case For $d_{x^2-y^2}$ Pairing in Cuprate Superconductors. *Physics Reports*, 250, 329–365.

Scerri, E. (2006) *The Periodic Table: Its Story and Its Significance*. New York: Oxford University Press.

Tinkham, M. (1996) *Introduction to Superconductivity*. New York: McGraw-Hill.

Tsai, W. and Kivelson S. (2006) Superconductivity in Inhomogeneous Hubbard Models. *Physical Review*, B73(21), 214510.

Waldram, J. (1996) *Superconductivity of Metals and Cuprates*. Bristol PA: Institute of Physics Publishing.

Zahar, E. (1989) *Einstein's Revolution: A Study in Heuristic*. La Salle, IL: Open Court.

AN INSTRUMENTAL BAYESIANISM MEETS THE HISTORY OF SCIENCE[1]

Commentary on "Modelling High-Temperature Superconductivity: Correspondence at Bay", by Stephan Hartmann

EDWARD JURKOWITZ

As this conference brought together historians and philosophers to examine a topic of common interest, in this response I critically examine Hartmann's Bayesian approach to analyzing science by studying how far it may be used to model real scientific fields and by locating areas requiring further philosophical work or generalization. Most broadly, Dr. Hartmann's paper seeks to find a consistent way to explain and characterize the continuity and coherence which the author finds to empirically characterize science in the past and present. In this response I first outline the program which Hartmann pursues and analyze the strategies used to pursue it. Next, while largely sympathetic to the project and underlining that it points towards ways of representing complex and partial forms of correspondence, I argue that the Bayesian formalism deployed is incapable of capturing the essential diversity of approaches and fluidity characteristic of theoretical work in real, living fields of scientific inquiry. It functions as a limited, post hoc formalism, and thus should be considered as one tool applicable for describing "normal science". It should be noted that, in this discussion, while I enter into the language of Bayesian discourse at various points, I also intentionally make statements which do not readily fit within the assumptions underlying Bayesian analysis, especially concerning the temporal characteristics of the proposed program.

In the first part of his insightful and clearly laid out paper Hartmann argues that the "Generalized Correspondence Principle", proposed by Heinz Post some years ago, represents a reasonable starting point for considering the sources of continuity in science. In examining the nature and limits of that approach to understanding continuity Hartmann takes the reader on a brief comparative tour of central positions which philosophers (and one historian of science) have taken on this issue.

[1] I thank Paul Teller for insightful comments and criticisms of an earlier version of parts of the paper on Bayesianism. All views expressed here are, of course, my own.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 131–141.

By critically studying the different kinds of correspondence relations relating successive scientific theories (Sect. 3), and through the study of one case of recent science, high-temperature superconductivity, he comes to the conclusion that, instead of a universal principle, the Generalized Correspondence Principle should be considered "at best, as *one tool* that might guide scientists in their theorizing."[2]

Hartmann proceeds to give a careful discussion of the different kinds of correspondence relationships which may link a predecessor theory (S) to a successor theory (L). He introduces basic set theoretic notions which will later serve him in introducing a Bayesian interpretation of correspondence, which he argues provides a fairly accurate account of the real relationships among theories in science, and a good account of the coherence characterizing science. After providing a set-theoretic formulation of Post's modified notion of the Generalized Correspondence Principle, according to which the "well-confirmed part of S [the old theory]" "is identical to the suitably translated part of L [the new theory]",[3] Hartmann recounts a range of objections which may be made against that claim. Many turn on variations of Kuhn's and related incommensurability arguments, which call into question the possibility of comparing the subparts of the old and new theories. The upshot of this discussion, presented in Sect. 3, is that while Post's description may have problems, and there may be a variety of different kinds of correspondence relationships which link predecessor and successor theories, when surveyed empirically, science reveals a significant amount of continuity. While the function of heuristic principles as *predictive* tools or guides for action is repeatedly questioned by the leading authors surveyed, Hartmann argues that, nevertheless, two kinds of correspondence, "Model Correspondence", and "Structure Correspondence" have "considerable heuristic potential" and "very high heuristic value" respectively.

This brings me to a first point: Hartmann's project is a double one. On the one hand, he seeks to frame a project which is capable of providing heuristic tools of use to working scientists. On the other hand, he seeks to provide an explanation of scientific change which is more realistic and more in line with the actual practice of science than alternative views. These dual goals should be kept in mind and separated when reading the paper.

Pursuing the first project, and drawing upon language and notions developed by philosophers such as Margaret Morrison, among others, Hartmann suggests that "the philosophical project of a methodology is best described by the picture of a *toolbox*." The methodologist's job is then to "abstract … a set of methods and techniques which can then tentatively be applied by practicing scientists in a particular situation." The different kinds of correspondence relationship he identifies may form part of this flexibly constituted toolbox, which scientists are then welcomed to pragmatically use in confronting the theoretical situations in which they find themselves. Hartmann's suggestion to categorize and tuck away in a toolbox the different forms of correspondence relation, so that scientists can then pragmatically use them to heuristically search for new theories related to current theories through different forms of correspondence, fits

[2] All in Hartmann, "Modeling High-temperature Superconductivity: Correspondence at Bay?," this volume, Sect. 1.

[3] This new theory may have a "condition", Q, imposed on it.

well with Hartmann's generally reflexive strategy of using methods and approaches found in science to inform and guide philosophical inquiry and practice.

With his methodologist's hat on, Hartmann argues, examining the research field of superconductivity in which little consensus has been achieved, that one must "trust the judgment of the scientific community" in deciding "the domain of applicability" of methodological principles such as the correspondence principles. More generally, since there are different kinds of continuity, Hartmann prudently advises against attempts to explain all kinds of continuity and theory change using some general form of correspondence relationship; to the contrary, since "general methodological principles unfold their fruitful action only when adapted to a specific problem", philosophers should restrict themselves to helping construct a toolbox of methodological principles from which scientists can draw as needed. To historians of science, for whom "practice" has become a keyword, and who regularly insist on the determinative force of local and even culturally specific forms of theoretical practice, this will likely be a welcome suggestion.

In Sect. 4 Hartman makes good on his program of learning from scientific practice by examining the case of theoretical work on high-temperature superconductivity. He surveys and contrasts more "conservative" strategies with Philip Anderson's "revolutionary" strategy to draw conclusions about the nature of correspondence and its role in science. Here he is faced with the question of what the Bardeen–Cooper–Schrieffer "theory" consists in; for one has to decide that before one can answer the question of what kinds of theoretical continuity exist in this field of theoretical physics research. However, what specifically Hartmann takes that theory to include is not fully specified – no doubt partly for reasons of space limitation. Although he distinguishes the "BCS theory" (basically pairing and Bose–Einstein condensation) from the "BCS model" (s-wave pairing), from the "BCS account" (the two together), this does not do the work of determining how far a new theory must deviate from a "core" of the old theory (here BCS) in order to be revolutionary, or even to be judged theoretically distinct from BCS. Any such judgment would require a further analysis of what the BCS theory is (and in particular asking whether it requires the BCS wavefunction and deciding what role the energy gap and related mechanisms play in defining the theory, as well as whether Bose condensation of the sort that BCS imagined is essential to the theory).[4]

More generally, deciding where exactly one (old) "theory" ends and a fundamentally distinct, or new one starts is essential to the Bayesian project outlined in the paper, but such decisions appear to be dependent on the temporal point at which they are asked, and on who is asked, as well as requiring an intimate understanding of the scientific fields and subfields in question. Doing the historical and analytical work of carefully characterizing the variety of approaches in a particular field at a given time, and detailing where they differ, are key prerequisites for accurately characterizing that realm of theoretical inquiry and determining the number of distinct competing theories, which, in turn, is a prerequisite for estimating the conditional probabilistic relationships assumed to link theories within the Bayesian program outlined later in the paper. From the historian's

[4] I avoid discussing any differences between my own interpretation of the relationships between the different theories of superconductivity and Hartmann's brief characterization of the field, as they are not central to the discussion of the issues at hand, and would require significant space.

perspective, the difficulties involved in deciding, amidst an intricate nexus of theoretical *approaches* and programs characteristic of active research, if and where theories divide from one another present important problems for this Bayesian project; for the Bayesian techniques of analysis are designed to analyze changes in degrees of confidence in sets of well-specified propositions subtending clearly identified theories. For example, one important difficulty arises when one theory does not appear, to most observers, to be essentially distinct from others at its inception, but in fact contains a core which, when worked out by cadres of theorists, yields a new synthetic perspective within that realm of theoretical activity. In such a case, while the starting method/approach/theory may be *retrospectively* recognized as a new theory, it is only from the *later* developed theoretical vantage point that the change-inducing starting point can be judged fundamentally new and distinct. Only in the altered theoretical landscape may it possibly enter into probabilistic calculations as a separately identified set of statements. Furthermore, comparing conditional probabilities referring to the previous viewpoint to those characterizing the newly framed theory appears historically artificial and practically difficult.

In the final section, "Rationality, Realism and Coherence", Hartmann briefly examines a range of explanations for the apparent continuity in science, ranging from forms of realism to "constructionism," in order to conclude that those and other explanations "do not explain the *relative weights* of continuities and discontinuities in scientific theorizing."[5] Again consciously deploying the strategies of scientists, Hartmann suggests that philosophers should "adopt a formal philosophical framework and construct *philosophical models* within this framework," and thereby follow scientific methodology by "constructing models of a theory."

He proposes to formalize a notion of "coherence", following an approach that he has begun to develop in his book with Bovens, *Bayesian Epistemology*, and to use that to describe putative aspects of generalized theoretical scenarios which could appear in a field of science (Bovens and Hartmann, 2003). In order to flesh out the basic idea of the coherence of a set of theories, first suggesting that "it is plausible that S (or better: S^*[the well confirmed part of the old theory]) and L [new theory] cohere more if, ceteris paribus, the two theories overlap more," he indicates a Bayesian representation of coherence. In Hartmann's project, detailed in his book, he argues that the "empirical part of … a scientific theory", T, is explicated "as a set of interrelated models", $\{M_1, \dots, M_n\}$, each of which is then represented by a set of propositions M_i which may itself be a conjunction of more elementary propositions. Hartmann then goes on to refer to a probability measure, P, over the set of model variables M_i. Next, after assigning probability relations between those variables depending on evidential and logical relations among the different models, he can assign a "coherence measure," coh(T), to a particular theory, T. This measure, he argues in detail in his book, can then be used to demonstrate things, to "explain" them. For example, if two theories share a set of models, then when the theories are taken together the sum is "more coherent" – given certain conditions.

However, it appears that such a program can be expected to work only where the previously noted difficulties concerning incommensurability, especially difficulties of translation between old and new theories, do not appear. For the whole formalism

[5] Hartmann, "Modeling High-temperature Superconductivity: Correspondence at Bay?", Sect. 6 (italics in source).

assumes that the competing theories are expressible in terms of a *common* set of propositions which underlie the models expressing the theory. (These are what the conditional probabilities are founded on, and what allows for comparison.) When incommensurability occurs, that assumption fails, and the claim to a well-specified "overlap" between theories is called into question in general. Moreover, the central Bayesian assumption of Hartmann's program, that the degree of confidence monotonically follows "coherence", appears questionable as well, since that assumption is premised upon the lower level assumption that how "expected the information is and how reliable the sources are does not vary from information set to information set", that the "expectance and reliability" are fixed, or in more general cases that they are heavily restricted.[6] But that condition could hardly be expected to hold when new results and theories are rapidly changing the expectations of theorists.

The author argues that a Bayesian account of the various correspondence relations discussed in Sect. 3 may help explain coherence. Shifting from the tool-box pragmatism used to prescribe methodological rules of action for scientists, this is an essentially post hoc project, which raises questions of historical accuracy and accessibility. A critical consideration of the capacities and successes of this sort of approach requires an examination of how the probabilities are to be assigned, and especially of *when* those are assigned. To wit: a retrospective Bayesian description may be valuable in allowing an analysis of the conditions which lead to continuity or discontinuity in particular scientific developments, and within theory change more generally, for it may allow a description of the patterns of conditional probability relations among models and theories which result in continuous, versus discontinuous, change within fields of scientific inquiry. However, from an historian's point of view, great difficulties surround any attempt to objectively assign probability relations in a post hoc manner, at least if the relations are intended to accurately describe the relationships *previously* existing among theoretical elements in a changing field of research. Assignment of probabilities is especially troublesome after there have been fundamental developments in a field. The difficulty of that project, however, casts doubt upon claims, like those of Hartmann, to have found a means of discovering general conditions and rules governing (empirically and/or normatively) the development of fields of inquiry. For, since the perspective on the field of research will have changed during the course of developments in the field, the subjective probabilities assigned post hoc to represent the connections between models (assigned by either scientific researchers or by Bayesian philosophers) are likely to differ from those that would have been assigned during the process of scientific development. If this is so, then any rules discovered by post hoc analysis would provide no warrant for the formulation of normative rules intended to guide scientists in the course of creative scientific research; for the set of conditional probabilities relevant to the contemporaneous situation may be quite distinct from the set of probabilities based on the retrospectively reconstructed view of the field. A rule found using one temporal stance does not necessarily imply a rule in the other. This issue is especially cogent since Hartmann (in a move I agree with as more historically accurate) allows contextual features into the estimations of the conditional probabilities connecting theoretical elements.

[6] Bovens and Hartmann 2003, Sect. 1.2.

Put more concretely, how can one say whether an *apparent* confirmation of the claim "coherence is truth conducive" using post hoc probabilities tells anything about whether that claim would be true with *contemporaneous* estimates of the probabilities? More generally, how can we know that the philosophical analyst does not assign the probabilities retrospectively in a way which skews the estimations toward confirming his preferred philosophical rule or theory?

Hartmann, within the context of his Bayesian formalism, shifts back to a methodological stance and makes the normative claim that scientists "*should* aim for highly coherent knowledge systems" because "coherence is truth-conducive." He further suggests that this can be proved within the formalism. Here he is making a mathematically true statement, but one which depends on a range of assumptions about the conditional probability relations connecting the relevant models and theories. Moreover, the difficulties involved in justifying the normative claim noted above are multiplied, since, within the Bayesian framework, Hartmann here is simply calling for a convergence of *subjective* estimates; but there is no guarantee that the claimed convergence is towards a true, rather than a false, theory, since it is only opinion that converges.[7] Indeed, contra the normative claim, it might be the case that, for example, greater theoretical diversity is conducive to greater long-term "truth." (Since more theoretical pathways are followed, this may increase the probability that a theoretical solution of greater stability is found). In that case theorists should not aim for greater coherence. Despite these problems, in theory, taking an unrestricted approach and applying a more general Bayesian analysis of the sort Hartmann proposes in an open, empirical fashion could allow the analyst to also discover this second result by characterizing different kinds of scientific scenarios. It must thus generate alternative normative claims (assuming one could avoid the aforementioned problems associated with the temporality of probabilities).

Any claim that a given set of probability relationships among theories, models, and propositions leads to increased coherence and confidence must clearly depend on a range of assumptions about the structure of the probability relations and their stability through time. Such assumptions would appear to exclude much of real, dynamic science. For at any point in time it may empirically happen that the conditional probabilistic relationships among models, and between evidence and models are *radically altered* by a new theoretical approach or viewpoint – something which has demonstrably happened many times in the history of physics (e.g., old quantum theory atomic models before, versus after, the advent of quantum mechanics). Put otherwise, if one aims to describe real sciences, at least dynamical, open ones, the probabilities appearing in the formalism cannot be taken as relatively fixed, or as having logically grounded unitary and separate relationships to the diverse pieces of evidence provided by experiment, but must be considered as (possibly radically) changing, mutually highly contingent values. Such variable probabilities run contrary to the conditions

7 Hartmann notes this in footnote 20, and Paul Teller has emphasized this elsewhere and to me.

needed to make Bayesian proofs of the sort Hartmann proposes relevant to reality.[8] Thus, if the Bayesian theorist really wants to claim that it is empirically – i.e., realistically and prospectively, not formally – true that "coherence is truth-conducive", then problems seem to arise.[9] It may be true, in accord with Hartmann's picture, that in certain situations there exist sets of reasonably assigned probability relations describing the subjective probabilities connecting evidence, models, and theories which have a structure such that the introduction of *any* shared model increases the coherence measure, *coh(T)*. However, it may also be the case, as Hartmann and Bovens note, that experts disagree significantly in evaluating the probability relations, which leads to a lack of a coherence measure and to indeterminism.[10] Moreover, even if the required conditions on the probability were satisfied at some time, say t_1, it is not clear that the increased coherence and confidence in a (set of) theory(s) generated by the introduction of an additional model would tell anything about the longevity of the theory(s), or about the course of empirically determined views of theorists at a later time, t_2. Why? Because theorists' estimates of prior and relative probabilities may shift radically when a new approach appears, a well-known problem for Bayesian accounts. In sum, since such proofs rely on restrictive assumptions about the constancy and structure of the probabilities involved, and the descriptions may only be formulated through retrospective reconstruction, the approach takes on a formal cast and leads one to question the extent to which it can capture the dynamical change characteristic of theorization in active physical sciences.

Just consider the probabilities for the case of high-temperature superconductivity. When assigning probabilities, who are we to ask and believe, Scalapino or Anderson? (Presumably they would give us different answers.) If slight modifications of the BCS theory represented by new pairing interactions are not yet a different theory, what about altered wave functions, or novel gap, statistical, and condensation effects? Moreover, whichever way one goes on whether a theory is distinct (and thus requires a separate set of values), what kind of values are we to assign the joint probability relations linking various theories, models, and approaches in order to describe the nature of the relationships connecting them? Asking the question highlights the implicit assumption

8 This approach assumes that in representing a scientific theory T "by a set of propositions $\{T_1, \ldots, T_m\}$", although the "constitutive propositions of a theory are tested in unison", nevertheless the propositions "are arranged into models that combine various propositions in the theory" in such a way that "each model M_i is being supported by some set of evidence E_i and that each M_i shields off the evidence E_i in support of the model from the other models in the theory and from other evidence. This is what it means for the models to be supported by independent evidence." Bovens and Hartmann 2003, Sect. 2.7 ff.

On the contrary, I think that most historians would insist that the probability values may have (and have had in a range of historical cases) variability way beyond anything that would be needed to fulfill such assumptions, or to make the broader claims connecting coherence with truth empirically valid. Faced with a contrary case, the Bayesian can, of course, say in response, "OK, the theorem does not apply in this case, since the conditions on the probability relations fail in this case." But how, then, can one independently test the hypothetical law (linking coherence and Bayesian truth), for the assignment of probabilities is subject to the discretion of the analyst?

9 Hartmann, "Modeling High-Temperature Superconductivity, …," in this volume, Sect. 6.

10 L. Bovens and S. Hartman (2003), Sect. 2.5.

that the probabilistic relationships are universally recognized rational relationships – rather than programmatic ones. Furthermore, at what point in time are we to evaluate those probabilities?

This set of issues suggests a broader question: what, exactly, is involved in the process of assigning the probabilities introduced in the formalism? The translation step which allows one to jump from possibly incommensurable old and new theories (which may contain quite disjoint theoretical estimations of the validity of particular arguments and models) to a set of "models," $\{M_i\}$, and a single probability function rating them, appears to flatten out – indeed squash – all the complexity that the historian sees in any given historically realistic theoretical situation. I doubt that the information-theoretic and semantic assumptions entangled with the claim to reduce working theories to sets of propositions which are to be inserted in a Bayesian calculus would pass muster with most historians (or with many scientists and philosophers) – certainly not when there may be instability or radical shifts in the theoretical structure of a field. Since Bayesian techniques were framed using concepts developed to determine degrees of confidence in matters of *fact*, this kind of approach seems unlikely to give a proper description of theoretical representations which are often programmatic, let alone of major theoretical shifts or paradigm changes in fields; for between such representations and across shifts there exist key divisions over the mathematical foundations and the logical structure of a (possibly nascent and as yet unspecified) future theoretical perspective or detailed theory. In periods of rapid change, contrary to a description like Hartmann's which "additively" assumes *more* coherence and more confidence to accrue to a cluster of theories whenever they share a model, it seems unlikely that one will even be able to assign a single probability measure.[11] More generally, if, on the one hand, the Bayesian project is a post hoc formalism which can only be used *retrospectively* to justify what in fact has occurred in a field, then, on the other, it can likely only be applied under conditions of relative theoretical stability in a field. However, while a link between Bayesian coherence measures and truth (pragmatic long-run truth) is problematic for reasons noted above, the empirical applicability of such a system of probability relations, one characterized by a cumulative structure, appears more promising for representing Kuhnian normal science.

Let me suggest (admittedly in a rough historian-of-physics fashion) a different, more flexible, and possibly more historically accurate application of a Bayesian-*like* formalism – all with the aim of pointing up the limitations of Bayesian descriptions for describing science in all its complexity. This will be a consistently contemporaneous description. Consider a kind of marketing-survey of a field of physics: one would survey the various theoreticians in a field about their views on the consistency and connections of each theoretical model and theoretical approach with

[11] To take an overly simple example for specificity, this would mean that under some conditions, since both classical mechanics and quantum mechanics use the harmonic oscillator model, the sum of those theories would be more coherent, an approach unlikely to accurately describe "instability" and shifts in a field of research. In contrast, in the description sketched below, assigning negative values to the "connections" among *multiple* competing, weighted "nets" or sets of relations describing a single field, may allow a better description of conflicted theoretical scenarios.

others. Perhaps one could even add in negative values, or "connection strengths", so that if a theorist viewed a model or theory as completely incompatible with the very foundations of the whole approach underlying the comparison model, the "connection strength" would be assigned some negative value, while a zero would reserved for the view that they were logically independent of one another. Having tallied up the responses from the surveys (possibly weighting the responses according to the respondent's seniority, and/or his university's research ranking), one would then likely have the best current estimate of the coherence relations among theories and models in the field as evaluated by the experts. With this – no doubt highly non-Bayesian – kind of network model of the field one could attain a lot more than a univalent coherence measure: one could get a measure of the "sociological-theoretical topography" of the field. There might be, for example, three largely distinct groups of researchers who produced similar estimations of the relationships linking different models and theories; but the three groups' estimations of the "connection-strengths" could differ significantly, even radically, one from another. Instead of *one* number characterizing the "coherence" of some theory or set of theories, perhaps such a sociologically more complex measure of the patterns characterizing researchers' theoretical views would yield a more complex and more accurate model of the relations among the theories and models in use at a given time within a field. By more "accurate" I mean allowing an analyst to more correctly determine correspondences between future theoretical developments (especially convergences) in the field and the multiple, *separate* webs or networks of researchers' estimated connection strengths which express the distinct outlooks of *different*, possibly competing, groups. Rather than settling for a single coherence measure, coh(T), a more complex set of relations between separate networks of researchers (and between their distinct connections patterns) and future theoretical developments could be tracked. Indeed, some move in this direction seems necessary since, as Hartmann admits, the experts in the field are likely the best estimators of the possibilities which different available theories afford for yielding workable explanations of the diverse phenomena under examination - and they rarely agree in highly active fields. A more complex historical model is especially necessary to describe historical periods of intensive research and controversy, where one competing approach may eventually win out and become dominant.[12]

I might add that in some ways this approach seems to be more in line with Hartmann's consistent efforts to pragmatically follow scientists when developing philosophical characterizations and models of science. Indeed, in theory such a non-Bayesian analyst of fields of theoretical work would be able to "empirically study" how, for example, a new set of discoveries, or new theoretical methods and models "shifted" the theoretical views of distinct, competing (possibly local) groupings of theorists in a given field – that by comparing surveys (patterns of connection-strengths)

[12] And hence only one of the competing networks of connection strengths (one groups' pattern) would turn out to empirically correlate with the future pattern.

before and after some set of events.[13] Such an approach would likely more closely follow the empirical modeling characteristic of much of science today.

This raises the more general question: What is it that the Bayesian framework gives one in this particular instantiation? Is it mainly that it provides a model of an essentially cumulative image of "coherentist" convergence of opinion, and possibly, implicitly, one tending toward "truth"? No doubt some will say it gives a description consistent with the "rationality" and continuity which they take to characterize the progress of science. But does such a Bayesian approach not implicitly tend to impose a restricted statistical model, possibly based on the inferential patterns of an idealized *individual* mind (representing an implicitly unified/unifying, or averaged scientific collective), onto a possibly more complex group of interacting theorists? Why not, then, consider a wider range of mathematical descriptions of scientific networks? The more complex sorts of description briefly indicated above would allow one to avoid attempting to explain away different "scientific tastes", or historically and sociologically very well-established differences between schools of research, especially by asserting, as Hartmann (and presumably other Bayesians) apparently finds necessary, when he states that "and yet, we [who?] observe that different scientists often assess new hypotheses similarly and reach similar conclusions about their acceptability". (Note that this very questionable claim is necessary in order to assign a unique Bayesian probability to each pair of models.) In short, the Bayesian approach should likely be taken as just one tool in an expanded philosopher's toolkit, one appropriate to limited conditions within the range of possible realistic historical theoretical scenarios.

While Bayesian descriptions may turn out to accurately describe certain conditions within a field of theoretical work, especially ones characterized by convergence of opinion, what if, for more general situations, we instead just give up the notion that there is some universal reference point from which the Bayesian analyst can accurately assign universal single probabilistic values to the connections between theories, models, and evidence? We could instead accept that different experts have essentially different estimations of those connections, and that those values may radically shift. Why not empirically model the different groups' viewpoints?

In short, I do not doubt that there are patterns to be found in the dynamics of theory change. I do doubt that the flattening out of "the empirical part of … a scientific theory" into a set of "propositions", along with the assignment – by the Bayesian philosopher – of a universal single-valued probability function across those, will allow the analyst of science to adequately capture the dynamics of theoretical work in an active field of science. As suggested above, it probably will not work very well if there are multiple schools with different competing approaches, or in rapidly changing fields of science. There one may have to consider models more general than those of this Bayesian framework.

[13] This approach would be better able to deal with complex theoretical situations in which conflicting approaches and sets of evidence arise. As opposed the Bayesian description Hartmann refers to, where an unstable, "saddle-point" theoretical situation might be described in situ as "more coherent", due to the averaging procedures involved in getting a single coherence measure/value, with such an alternative approach one could understand the instability as resulting from competition and frustration between competing networks.

In the end, probably like most historians of science, I find the work of delineating the nature of the correspondence relationships that have emerged in the historical development of science very productive, and Hartmann's general project of modeling the theoretical organization of fields of inquiry a tantalizing prospect, one promising to yield a more complex philosophical description of theoretical work in various fields, but I find the Bayesian formalism deployed and its underlying assumptions too thin to model all of the complexity and diversity present in science as it actually and historically develops. In spite of these limitations, this paper is to be commended for pointing the way towards pragmatic approaches to modeling science which may ultimately capture both its coherence and its complexity.

BIBLIOGRAPHY

Bovens, L. and Hartmann, S. (2003) *Bayesian Epistemology*. Oxford: Oxford University Press.

PART 5

FROM THE CUMULATIVITY OF PHYSICAL PREDICTIONS TO THE CUMULATIVITY OF PHYSICS

IS SCIENCE CUMULATIVE? A PHYSICIST VIEWPOINT

BERNARD D'ESPAGNAT

Abstract In view of the facts that when a recipe has been found working it will continue to work and that, in physics, the equations of superseded theories usually remain useful, the statement made by some epistemologists that science is not cumulative is difficult to understand. Some might try to relate its validity to a "principle" according to which, to be "really scientific" a theory must be such that the set of its basic laws does not explicitly refer to us and the predictions we issue. But it will be shown that if this principle is consistently adhered to quantum physics must be taken to be a non-scientific theory. Since this conclusion is absurd the "principle" should be rejected, and science viewed as cumulative.

Keywords Cumulativity, Kuhn, Poincaré, Quantum Physics.

I. It was a great stir when Kuhn and Feyerabend explained that science mainly develops through sequences of revolutions – "received" theories being then replaced by other ones grounded on different notions – and when they, or their followers, inferred from this that science is not cumulative.[1] In philosophical circles as well as in the general public this was taken to be a truly ground-shaking new idea. But, was it really quite as new as was then believed? It seems appropriate to have a look at what the mathematician Henri Poincaré – who was also a keen philosopher – wrote on the subject. The following is a sentence taken from his book *La science et l'hypothèse*:

> Men of the world wonder how short-lived our scientific theories actually are. After but a few years prosperity they see them being dropped, one after the other. Ruins, they observe, accumulate upon ruins. They foresee that presently fashionable theories will soon meet a similar fate, from which they infer that none of them is trustworthy. It is what they call the failure of science.[2]

[1] Is this claim really theirs? In Kuhn's introduction to *The Structure of Scientific Revolutions* there are sentences that strongly speak in favor of the answer "yes". But anyhow, what we are here interested in is not so much what Kuhn really thought and stated (this is a matter for Kuhn experts) as the relevance of the views on science that, rightly or wrongly, many of his followers as well as the general public derived from his writings. The (devastating!) idea that science is not cumulative is undoubtedly one of them.

[2] "Les gens du monde sont frappés de voir combien les théories scientifiques sont éphémères. Après quelques années de prospérité ils les voient successivement abandonnées; ils voient les ruines s'accumuler sur les ruines; ils prévoient que les théories à la mode devront succomber à leur tour à bref délai et ils en concluent qu'elles sont absolument vaines. C'est ce qu'ils appellent la faillite de la science". (La science et l'hypothèse, Paris, 1902, chapitre 10)

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 145–151.

Except for the fact that these lines were written in year 1902 – more than a century ago! – you might think that the "men of the world" whose ideas they so keenly describe are Kuhn and Feyerabend's disciples. For indeed if I may try and summarize what Kuhn stated – or at least forcefully suggested – I think it goes like this. During a transition period, it often happens that two groups of scientists uphold paradigms that are incompatible with one another. Hence in the substitution, which finally takes place, of one of them by the other one a historical, therefore contingent, element comes in. A partly consistent conception gets replaced by another one, which also is but a partly consistent one. In other words, one group gained the victory over the other one.

Faced with such reasoning we may well wonder whether, by any chance, science might not just be a social fact, somewhat comparable to tastes or fashions. And indeed quite an appreciable number of contemporary thinkers share such a view, in many countries and in France in particular. As an example let me just put forward a quotation from Dominique Pestre. Writing about the views of Bloor and Collins, this contemporary French epistemologist wrote:

> They stress that […] every piece of knowledge should be considered imbedded in, and intermingled with, anthropological situations *that are constitutive of them* [emphasis mine][3]

and he expressed his approval of this standpoint. So, in short, we are now in a situation in which, schematically, the views of the "men of the world" Poincaré referred to are taken up by learned epistemologists who back them with elaborate arguments. On the other hand, Poincaré himself was very far from sharing the views in question. In fact he mentioned them just to explain why he did not share them. I suggest we should have a look at his reasoning. He wrote:

> No theory seemed more firmly grounded than Fresnel's one, which attributed light to motions of aether. However Maxwell's one is now preferred. Does this mean that Fresnel's theory was of no value? No, for Fresnel's purpose was not to know whether or not aether really exists, whether or not it is constituted of atoms and whether these atoms really move in this or that way. It was to correctly predict the optical phenomena. Now, Fresnel's theory still allows for such predictions, today just as before Maxwell. Its differential equations remain true, they can be integrated by the same procedures as before, and the outcomes from these integrations still preserve their full value.[4]

Indeed, equations remain valid (at least approximately) and observational predictions remain correct. And this is true concerning, not only Fresnel's theory but practically all now "superseded" great theories. These are two facts that are, I think, of quite crucial significance. And, to say the truth, I am very much surprised when I observe

[3] "Ils insistent pour que […] nous pensions les savoirs comme toujours inscrits et imbriqués dans des situations anthropologiques qui leur sont constitutives". (Pestre, 1997)

[4] "Nulle théorie ne semblait plus solide que celle de Fresnel qui attribuait la lumière aux mouvements de l'éther. Cependant on lui préfère maintenant celle de Maxwell. Cela veut-il dire que celle de Fresnel a été vaine? Non car le but de Fresnel n'était pas de savoir s'il y a réellement un éther, s'il est ou non formé d'atomes, si ces atomes se meuvent réellement dans tel ou tel sens; c'était de prévoir les phénomènes optiques. Or cela, la théorie de Fresnel le permet toujours, aujourd'hui aussi bien qu'avant Maxwell. Les équations différentielles sont toujours vraies; on peut toujours les intégrer par les mêmes procédés et les résultats de cette intégration conservent toujours toute leur valeur". (Poincaré, 1902, Chap. 10)

that most epistemologists just superbly ignore them. Since we just spoke of Maxwell's theory, let us take Maxwell's equations as an example. Think of how many times their tentative interpretation changed, from the year they were discovered to the present. According to the standpoint of Poincaré's "men of the world" (and many of Kuhn's followers) all these changes should count as so many crumbling down of theories, compelling scientists to start on totally new bases. But this is just not so! Maxwell's equations – which are the real substance of Maxwell's theory as Hertz pointed out – are still fully and totally "with us". This is obvious as far as technology is concerned. The existence of portable phone receivers and Martian robots definitely convinces (or at least should convince!) engineers and executives that Maxwell's equations did not just reflect temporary social prejudices. And the same holds true concerning pure science. When dealing with laser beams and other quantum electrodynamics phenomena, present-day theorists do make use of the equations in question and would indeed be completely lost if they were deprived of them.

Hence, to put it in a nutshell, Science constitutes a remarkably synthetic tool enabling us to predict what we shall observe and, as time goes by, this predicting power steadily increases. Since everybody, Kuhnians included, agree with this, everybody must agree with the inescapable conclusion that, in this respect, science is cumulative.

Unfortunately it seems that neither Kuhn nor his disciples discussed the problem from this angle. True, competent "Kuhnians" pointed out to me that, although Kuhn granted that science is, instrumentally viewed, a progressive enterprise with true cumulation of results, still "this was not his point". Clearly, however, such an answer is here too short. You cannot, at the same time, claim, which is the thesis here under debate, that "science is not cumulative, period", and grant that it is cumulative in a domain you are not interested in. And hence, in order to salvage the thesis in question its defenders have no other choice than to rely on the presence, in the foregoing conclusion, of the restrictive words "in this respect". They have to convince us that this cumulative production of predictive tools, somehow, "doesn't really count", or "is not science proper". At least, this is quite obviously what they should do.

II. In fact, however, I am not aware that any of them took up this line of reasoning. In a sense it is a pity since the burden of the proof that science is not cumulative clearly falls on them. On the other hand it is interesting to try and fill up this blank, just in order to see whether or not it is at all possible to preserve the "science is not cumulative" thesis by proceeding as I just said, and this is what we shall presently do. It must, of course be understood that, correlatively, the arguments we shall consider and critically analyze in this second part do not emanate from the Kuhnian school. They are, to repeat, just arguments that defenders of the no cumulation thesis might conceivably imagine putting forward.

Clearly, to repeat, the only way in which one may conceivably hope to attain this objective consists in restricting the definition of what we call "science" in such a way that theories should be considered scientific only if there is more in them than just prediction of observations. One such restriction comes quite naturally to mind. This is the view that what is requested is explanation, that some laws do explain and some do not and that, in particular, laws that merely predict observational results do not explain. Indeed, many scientists (see, e.g., Bunge, 1990, p. 40, p. 97) have been stressing for a long time the appropriateness of such a specification. The difficulty with it is that the notion of explanation is a delicate one. It is generally considered that to explain consists in connecting up a great many different facts by showing that they come under the same set of basic laws. But for our purpose we must not, of course, include observational predictive laws in this set (since otherwise the no cumulation thesis would immediately break down). Finally therefore, to attain our objective it seems natural to require that the laws composing the theory be expressed by means of statements in which no explicit reference is made to us, that is, to what we might do and observe as consequences thereof. For future reference let us call this requirement "requirement R".

Admittedly, the theories that most straightforwardly fulfill requirement R are the "realist" ones, in the philosophical sense of the word. They are the theories that are descriptive, in the sense that the basic concepts they make use of satisfy the two conditions of being in correspondence with our experience and being "ontologically interpretable". Unfortunately, these conditions nowadays appear to be unduly restrictive. The reason is that standard quantum mechanics does not satisfy the second one and that, in view of non-separability and so on it seems extremely doubtful (to say the least) that any theory susceptible of replacing standard quantum mechanics could ever satisfy it. Now quantum mechanics, as it is well known, is not just the only theory that accounts for our experience in the so-called microscopic realm. It lies at the basis of, not only chemistry and nuclear science but also quantum optics (lasers), solid-state theory, statistical mechanics, high-energy physics and so on. Unquestionably it is a scientific theory. Any criterion grounded on ontological interpretability would set it aside and is therefore unacceptable.

It follows that the "science is not cumulative" thesis cannot be salvaged by restricting admissible theories to "realist" ones. We must try and find a milder condition. A priori it seems that this should be feasible for it is well known that science is reconcilable with antirealism. The basic laws of, say, a physical theory are composed of statements bearing on physical quantities and in antirealism a statement is not defined to be true on the ground that it describes reality "as it really is". According to Dummett, in antirealism the truth of a statement is essentially linked to human knowledge. More precisely (in Dummett's words): "a statement […], if true at all, can be true only in virtue of something of which we could know and which we should count as evidence for its truth" (Dummett, 1978, p. 145).

Let us then investigate in some detail to what extent it is possible to build up theories whose basic statements are of this type and in which no explicit reference is made to us. This investigation we shall first carry through within the realm of classical physics, and we shall see that indeed in that realm it is possible to do so. We shall then examine whether or not this is the case concerning quantum physics.

To begin with, let us then consider classical physics. Applied to the notion "value of a physical quantity" Dummett's above-stated criterion becomes: "For a statement concerning the value of a physical quantity to have a definite truth-value (be true or false) it is necessary that the physical quantity in question be measurable (either directly or indirectly)". Fortunately, in classical physics it turns out that this criterion is also a sufficient one. For, within the classical realm, we consider it obvious that any quantity that can be measured ipso facto has a value, quite independently of whether or not we actually measure, or shall measure, it.

Note moreover that, within classical physics, the said criterion is applicable quite independently of circumstances. Any measurable quantity is considered to have, at any time, a definite, known or unknown, value. This has the important consequence that, within classical physics, laws describing the evolution of systems normally ("automatically" we might even be tempted to say) – satisfy Requirement R. Indeed these laws describe how quantities pertaining to the observed system evolve in time and since we know (from continuing measurability) that when these quantities change they keep on having definite values, there is no reason that the accounts these laws yield of the said changes should mention us. And in fact, within classical physics as it is presently constructed, they do not (except within some formulations of statistical mechanics, but this is another subject).

Hence, as we see, within the realm of classical physics "antirealist" theories may be constructed that are not centered on the premiss of ontological interpretability but still satisfy Requirement R. Now the question is: does the same hold true within the realm of quantum physics?

Concerning this question the first point to be noted is that, contrary to what is the case in classical physics, in quantum physics, measurability is not a sufficient criterion guaranteeing that a physical quantity has, objectively, a value. Imagine for example that we are interested in the precise position, at a given time, of an electron that has been accelerated and was imparted thereby a given momentum. We do dispose, in principle, of instruments enabling us to measure its position as accurately as we like, and this is why, in quantum physics, it is asserted that position is a measurable quantity. Still, we cannot infer just from this (from this possibility we have of measuring it) that the proposition "at such and such a time the position coordinates of the electron have such and such values" has a truth-value (is true or false). This is due to the fact that, in such cases, the measurement in question, if performed, merely informs us of the values these coordinates have after it has been made. For indeed, in general this measurement has – depending on the theory or model we believe in – either created or modified the said values. It did not just simply register them.

Now, does this fact quite generally prevent us from ascribing definite values to physical quantities? And from producing, this way, useful descriptions of quantum systems? No, its consequences are not so drastic. For example, in the case of the just considered electron we may conventionally describe it as having quite a definite momentum. Similarly, atoms may often be said to be in such and such energy states, etc. But still, it can never be the case that this possibility be independent of the circumstances in which the system in question actually is (or "was prepared"). There are many circumstances in which electrons have no definite momenta, and there are some in which atoms have no definite energy. And, what is more, it may quite well happen that in the course of time, due, for example, to an encounter with other systems, a physical system evolves in such a way that while initially it was endowed with a physical quantity having a well-defined value, in the final state this is no more the case. It follows from this that, contrary to what we have seen is the case in classical physics, even within an "antirealist" approach the set of the general quantum mechanical evolution laws cannot be considered to yield a description of how physical quantities endowed at some time with definite values evolve in time. Indeed it cannot be imparted a form satisfying Requirement R. I mean: to apply in a meaningful way the said set of laws we must, at some stage, explicitly make use of some elements in it that refer to us and what we shall observe.[5] In other words we must give up the view (assuming we entertained it) that switching to antirealism as quantum mechanics requires would ipso facto remove the difficulty we are in. We, willy-nilly, have to take prediction of observations into account in a basic way, which implies that in quantum physics the cumulative character of the observational predictive laws may not be set apart from the set of the elements that serve to characterize science proper. (It is true that the law describing the evolution of wave functions in-between successive measurements makes no reference to us, but this does not constitute a counter-example to what I just said since this law is just an element of the set in question and is such that considering its evolution independently of the other elements makes no sense).

III. Let us recapitulate what we did. In part I of this text we noted that the twentieth century last decades witnessed the wide spreading of an idea said, rightly or wrongly, to have been inspired by Kuhn, namely the view that science is not cumulative. This we contrasted with an older conception of Henri Poincaré who considered its ability at successfully predicting new observational results to be one of the most significant characteristics of science, noted that, as time goes by, this ability steadily increases and concluded that science is cumulative. With some puzzlement we then observed that, apparently, Kuhn himself and his followers agreed that science is cumulative in this respect, while continuing to claim that science is not cumulative. Not having found in the Kuhnian literature a solution to this apparent inconsistency, we tried to invent one ourselves. And what we did essentially amounted to observing that to this end it is necessary, when stipulating the conditions a discipline must satisfy in order to be called a science, to require that, at any rate, observational predictive power should not be the central element of the discipline in question.[6] This is a very mild requirement and we tried to take advantage of the fact that in most customary conceptions of science it is fulfilled, since, in them, what is considered essential in science is primarily its power at describing and explaining. We found that in the abstract such a way of characterizing science does indeed remove the difficulty but that it fails to do so when specific features of the sciences as they actually are taken into account since the observational predictive power does constitute the core of one most important discipline, namely quantum physics.

I do not claim that the above exploration is exhaustive. It is conceivable that there should exist other characterizations of science, both restrictive enough to leave out its – obviously cumulative – predictive power and general enough to accommodate quantum mechanics. But this conjecture is doubtful for, as we saw, quantum mechanics

[5] As a rule this takes place though the channel of probabilities that are of a subjective – or, better to say, inter-subjective – nature, and the main one of the elements just alluded to is called the (generalized) Born rule.

[6] Since, when it is, it is obviously impossible to consider the cumulativity of the "recipes" this power generates to only be a side aspect that "doesn't count".

comes as a tardy defense of Poincaré's views: what is most secure in it is, by far, the set of its observational predictive rules.[7]

Anyhow the burden of finding a way to remove the difficulty clearly falls – to repeat – on the supporters of the claim that science is not cumulative. As long as they do not produce one we shall be fully entitled to judge that science is cumulative.

Acknowledgements

It is a pleasure for me to express here my hearty thanks to Léna Soler for highly helpful discussions.

BIBLIOGRAPHY

Bunge, M. (1990) Des bons et des mauvais usages de la philosophie. *In L'enseignement philosophique*, 40(3), 97–110.

Dummett, M. (1978) *Truth and Other Enigma*. London: Duckworth.

Pestre, D. (1997) *Histoire des sciences, histoire des techniques*, EHESS, Paris.

Poincaré, H. (1902) *La science et hypothèse*. Paris.

Worrall, J. (1989) Structural Realism: The Best of Both Worlds? *Dialectica*, 43, 99–124.

[7] In fact, as John Worrall pertinently stressed (Worrall 1989), in his above-quoted book Poincaré continued his defense of cumulativity by pointing out that, although the *nature* of light was quite different in Maxwell's theory from what it was in Fresnel's, lts mathematical *structure* was just the same. Even though quantum physics is predictive rather than descriptive this extended argument may also be carried over to it, at least in a milder form. We do not speak any more of "oscillations of *something*" but we still use such words as "frequency", "wave-length", etc., and the mathematical structure of the quantum electrodynamical Hamiltonians are directly inherited from Maxwell's electrodynamics. Let me express my thanks to the anonymous referee who reminded me of the relevancy of Worrall's article to such matters.

COMMENTARY ON "IS SCIENCE CUMULATIVE? A PHYSICIST VIEWPOINT", BY BERNARD D'ESPAGNAT

MARCEL WEBER

Bernard d'Espagnat's concern is to meet some of the challenges of Kuhnian philosophy of science, especially as they concern the claim that scientific development is not cumulative. His main line of argument is to use the fact that, in physics, the equations of superseded theories frequently remain useful, for example, for making predictions. How can this be reconciled with the thesis of non-cumulativity? D'Espagnat argues that it cannot and that, therefore, this thesis ought to be rejected. In my commentary, I want to argue that d'Espagnat's arguments fail to refute Kuhn's claims with respect to cumulativity. Thus, I will defend Kuhn's claims only against the arguments provided by d'Espagnat; I do not want to claim that Kuhn's position is the ultimate truth.

I want to begin by briefly summarizing Kuhn's (1970) own position as regards cumulative scientific change. The first thing to note is that Kuhn's thesis of non-cumulativity pertains only to *revolutionary* scientific change. Only scientific revolutions bring about the radical changes in the cognitive fabric of scientific theories that give rise to non-cumulativity. By contrast, normal science is fully cumulative. Normal science is likened by Kuhn to the practice of puzzle-solving, where the puzzles to be solved as well as the solutions produced stand in an appropriate similarity relation to some paradigmatic problem solutions. In fact, the very nature of normal science makes normal scientific development cumulative, as each newly solved puzzle is simply added to the stock of puzzles already solved and the latter are in no way affected by the coming into being of new puzzle solutions.

What about revolutionary scientific change? Here, the situation is altogether different. Scientific revolutions put the paradigmatic problem solutions of a scientific tradition out of service and replace them by new problem solutions that take over the role of guiding normal scientific research. Famously, Kuhn has argued that the theories that are part of a new paradigm are incommensurable with those of superseded paradigms. This does not mean that they cannot be compared, only that a point-by-point comparison is impossible. "Point by point" means that, for each prediction of the newer theory, there is a corresponding prediction of the older theory to which it can be compared for their match with empirical data. However, if theories are incommensurable in Kuhn's sense, then one theory makes predictions that have no match in the other theory. One theory may also be more accurate in a part of its domain, while the other theory may be more successful in other parts, while there is no way of weighing these merits and disadvantages. Such as situation, according to Kuhn, makes it impossible

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 153–156.

for an individual scientist to make a rational choice between the two theories. Therefore, Kuhn compared the adoption of a new paradigm by an individual scientist to a "conversion experience." However, this does not imply that the development of science through scientific revolutions *as a whole* is irrational or that it is determined by social factors alone. Because if an old paradigm really ceases to function and consistently fails to solve certain intractable puzzles, more and more scientists will adopt the new paradigm. Therefore, there is some sort of emergent rationality in the scientific community. Furthermore, this paradigm shift as a whole is not "contingent," as Professor d'Espagnat understands Kuhn. Far from it: A new paradigm picks up where the old paradigm failed, and the failure of the older paradigm as well as the success of the new are a result of the fit of theory and nature. The only thing that is contingent in Kuhn's picture is the choice of a theoretical framework or paradigm by the individual scientist.

What exactly does non-cumulativity mean in the context of Kuhnian revolutionary change? As already indicated, it is a major difference between scientific change in the normal and revolutionary phases that, in the former phase, new problem solutions are simply added to the existing stock in accordance with the similarity relations to the paradigmatic problem solutions, while in the latter the relevant exemplary problem solutions are replaced by new ones. It is precisely here where Kuhnian scientific change exhibits its non-cumulative nature: the new paradigmatic problem solutions are not simply additions to a pre-existing stock of problem solutions; they rather *replace* the older solutions. This replacement can bring about a radical reinterpretation of the older problem solutions. To give an example, a falling stone is interpreted very differently in Newtonian mechanics and in General Relativity Theory. In the former view, it is interpreted as a mass that is attracted by a force, while according to the latter it is a force-free motion along a geodesic in curved spacetime. According to Kuhn, it is not possible to intertranslate the two descriptions of this basic event, as the central concepts of the two theories are incommensurable. Again, this does not preclude us from saying that there is a sense in which the relativistic description is superior, but this superiority does not consist in its better explaining the empirical facts as stated in a neutral observation language (as older philosophies of science would have it). It rather consists in the fact that general relativity theory was able to solve cases where Newtonian mechanics failed (such as the perihelion of Mercury).

What is the fate of some of the established problem solutions of an older paradigm after that paradigm has been replaced? This is a difficult and subtle issue and offers plenty of room for disagreement between physicists and Kuhnians. For example, there are the much-discussed "limiting case" examples. Special relativity turns into Newtonian physics when velocities are much smaller than the speed of light (or, alternatively, when the speed of light approaches infinity). How can this be squared with the thesis of non-cumulativity? The Kuhnian's answer is straightforward: While it is correct that the *predictions* converge in these limiting case scenarios, the relevant *ontologies* do not. It is still the case that mass is a relational property in relativistic physics, while it is an intrinsic property in classical mechanics. According to Kuhn (and also P. Feyerabend), the meaning of such theoretical concepts changes radically in scientific revolutions, and the way in which the phenomena are interpreted changes with them. This is the sense in which revolutionary change is non-cumulative.

Now what about the phenomenon that, in some cases, the equations used to formulate scientific theories are retained in scientific revolutions? First of all, it should be noted that Kuhn was well aware of this phenomenon. According to Kuhn, the role of equations is that of "symbolic generalizations" or "law-schemata" or "law-sketches."[1] These schemata contain placeholders instead of empirical concepts (e.g., the letters *F*, *m* and *a* instead of the interpreted concepts of force, mass and acceleration). Any kind of physically adequate description of a situation must begin with such a law-schema. The tricky part is to learn to see a new problem situation as an instance of the law schema. It is in this step that the whole cognitive apparatus of a scientific theory (as contained in the family resemblances in the exemplary problem solution) is contained. The equations are mere formulas from which everything but the logical and mathematical concepts (i.e., concepts like equality, multiplication, integration, derivative, etc.) have been abstracted away. It is for this reason that Kuhnians are not impressed by the retention of equations across revolutionary divides. Even though this may be seen even under a Kuhnian framework as a curios fact that requires explanation, it does in no way affect the thesis of non-cumulativity. For according to Kuhn, everything *but* the logical and mathematic concepts have changed in the revolution.

To illustrate this on d'Espagnat's example, Maxwell's equations as such are merely symbolic generalizations. They have to be *interpreted* by assigning one of their theoretical magnitudes the significance of expressing the strength of an electromagnetic field. However, this interpretation is subject to scientific change. For example, electromagnetic waves as they are treated by Maxwell's equations were once thought to travel in a luminiferous ether. This, of course, is no longer believed by contemporary physicists. Therefore, the descriptive content of electromagnetic theory has been subject to replacement. In this example, scientific change was not cumulative.

To this objection, d'Espagnat addresses the following counter-objection: The objection assumes certain standards of interpreting physical equations. However, these standards are not satisfied by one of the most fundamental physical theories that there is: quantum mechanics. The equations of quantum mechanics cannot be interpreted in the way that, for example, Maxwell's equations were once interpreted (e.g., by assuming a luminiferous ether). Therefore, the standards of interpretation must be dropped altogether. Physical equations have no other purpose than to *predict* possible experiences, not to describe reality. There is no interpretation. But many successful predictions of scientific theories are, in fact, retained across scientific revolutions. Hence, scientific change is cumulative after all.

I think that such a reply begs the question against the Kuhnian thesis of non-cumulativity. For Kuhn argues at great length how the (phenomenal) world as it appears to the scientist changes radically in a scientific revolution. The phenomena simply are not the same any more, even if it should be the case that some numerical values predicted by an older and a newer theory are the same. To claim that some predictions are retained during scientific revolutions amounts to flatly denying what Kuhn argues for. Of course, it is possible to deny this, but this requires an argument. I cannot

[1] For the following, see Hoyningen-Huene (1993, pp. 103–104).

find such an argument that goes beyond simply assertion in d'Espagnat's text. This is why Bernard d'Espagnat, while raising many interesting and important issues, has not really addressed the Kuhnian challenge to cumulative scientific change.

BIBLIOGRAPHY

Hoyningen-Huene, P. (1993) *Reconstructing Scientific Revolutions. The Philosophy of Science of Thomas S. Kuhn*. Chicago, IL: University of Chicago Press.

Kuhn, T. S. (1970) *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

PART 6

FROM DENOTATIONAL CONTINUITY TO ENTITY REALISM

THE OPTIMISTIC META-INDUCTION AND ONTOLOGICAL CONTINUITY: THE CASE OF THE ELECTRON

ROBERT NOLA

> *The first observer to leave any record of what are now known as the Cathode Rays [subsequently renamed 'electrons'] seems to have been Plücker.* (J. J. Thomson, May 1897, p. 104)

Abstract The pessimistic meta-induction attempts to make a case for the lack of ontological continuity with theory change; in contrast, its rival the optimistic meta-induction makes a case for considerable ontological continuity. The optimistic meta-induction is argued for in the case of the origin, and continuity, of our talk of electrons (even though the term "electron" was not initially used). The case is made by setting the history of identifying reference to electrons in the context of a generalised version of Russell's theory of descriptions, Ramsey's theory of theoretical terms and a development of these ideas by David Lewis.

Keywords Pessimistic Meta-Induction, Electron, Scientific Realism, Ramsey Sentence, Russell's Theory of Descriptions, Ontological Continuity and Theory Change.

When philosophers wish to cite an example of a theoretical entity whose existence has been established by science, more often than not they cite the electron. Scientists and most philosophers of science tend to be realist with respect to electrons; that is, they think that electrons exist in a strong sense independently of any minds or any theories or languages minds might use to talk about electrons. Millikan even thought that he could see electrons, telling us in his 1924 Nobel Prize lecture:

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 159–202.

> He who has seen the [oil-drop] experiment, and hundreds of investigators have observed it, has literally seen the electron.… But the electron itself, which man has measured … is neither an uncertainty nor an hypothesis. It is a new experimental fact that this generation in which we live has for the first time seen, but which anyone who wills may henceforth see. (Millikan, 1965, pp. 58–59; emphasis in the original)

Realists might demur from such a strong perceptual claim, saying that even if Millikan had an electron directly in front of him, he did not have an electron as an item in his field of vision since they are too small to be seen by our eyes. At best he saw some characteristic movements of oil drops in an electric field and inferred that electrons were present as they hopped on or off the oil-drops causing them to move up or down. But at least Millikan's position connotes a strong realism about electrons, as strong as that we suppose about the ordinary objects we see.

Such a robust realism has been questioned in a number of different ways; the focus here is on just one, the Pessimistic Meta-Induction (PMI). The argument comes in a number of different forms. In Sect. 1 the version first given currency by Putnam is discussed. The argument raises the possibility that Rutherford, Bohr, later Quantum theorists, etc. were not talking about the very same entity, the electron, owing to the very different theories they held about electrons. It will be argued here that PMI should not be accepted; in its place OMI, the Optimistic Meta-Induction, will be advocated. In the case of the electron this means that despite radically changing theories, scientists did talk about the very same thing, electrons, even though they did not use the term "electron" for a considerable time. How is this to be shown? Section 2 sets out a version of Russell's theory of descriptions generalised to apply not only to individuals but to kinds like the electron. It is shown that Russell's theory is a special case of the Ramsey Sentence as developed by David Lewis. This background semantic theory is employed in subsequent sections to show how the same entity can be identified despite changes in theory of the entity so identified. It will emerge later that such descriptions carry the burden of fixing a denotation (despite the quite different names used to denote the kind of entity so identified) in contrast to the full Ramsey Sentence which shows how reference can fail.

Section 3 considers Plücker's 1857 discovery of an unusual phenomenon and the postulation of a "something" as cause of it. Section 4 considers the growth in our knowledge of the "something" and some of the names that were introduced to refer to it, "cathode rays" being the name most widely adopted. Section 5 considers how different theories about the very nature of the "something" were proposed, from wave to particle, while there was continuing reference to the same "something". This requires an account of how the Ramsey-Lewis procedure can be used to identify the same "something" without invoking the full theoretical context in which a term occurs. Section 6 considers Thomson's two important papers of 1897, his summary of relevant experimental knowledge of the time and his unusual theory about the "something". It is in this context that the difference between the use of Russellian descriptions and the full Ramsey sentence becomes important; the former enables ontological continuity to be established while the latter shows how much of what Thomson wanted to say about his "corpuscles" could not apply to any entity. Section 7 makes some brief comments on how the term "electron" was introduced into science in multiple ways; but what is important here is the dependence of each introduction on already well-established definite descriptions. The term "electron" might be multiply ambiguous but this carries no threat of radical incommensurability in science.

Section 8 makes a link to the historical story told in Bain and Norton (2001) which begins with the classical theory of the electron and continues it up to various current quantum theories of the electron. The semantic issues raised by the Lewis version of the Ramsey Sentence come to play an important role in forging this link.

When historians of science come to talk about some episode in the history of science, such as the development of our views on electrons, they often speak of the *concept* of an electron (either with or without relativisation to the individual who entertains the concept, a period of time, etc.). However, the concept of a concept is a notoriously difficult one in philosophy, and given our unclear understanding, it might not be able to discharge the burden that is placed on it, especially in its use to analyse what is going on in science. One difficulty arises in talk of conceptual change. If a concept undergoes change (whatever this might mean), just how much change can it undergo and remain the same concept, and just how much change leads to a different concept? Historians of science get caught up in such matters concerning conceptual change. However, it would be much better to avoid them, and they are avoided here. A different approach would be to eschew concepts in favour of talk of sets of beliefs held by a person at a time. Then one can trace how the sets of beliefs differ from person to person, or from time to time. Here the flourishing theories of belief revision might be of better service than talk of concepts. Turning to a different matter, there should also be a focus on what the concepts are about rather than the concept itself. It remains notoriously difficult to say what the extension of a concept is and whether or not the extension changes with change in the concept. A way of avoiding this problem is suggested in Sect. 2 of the paper, and is then adopted throughout to show that despite radical changes in our "concept" of the electron it is the same thing that the different concepts are about. The traditional theory of concepts cannot really deal with this problem, but the development of the Ramsey sentence by Lewis outlined in the paper can accomplish this. Though we cannot eliminate all mention of concepts, problems that they generate for continuity of ontology can be by-passed.

1. THE PESSIMISTIC META-INDUCTION VERSUS THE OPTIMISTIC META-INDUCTION

Putnam expresses one form of the Pessimistic Meta-Induction, PMI, in the context of discussing Kuhnian and Feyerabendian incommensurability, illustrating it with the (alleged) incommensurability of the early Bohr-Rutherford theory of the electron compared with the later Bohr's theory of the electron, and even our current theory of the electron. Is it the same item, the electron, which is being referred to in all of these theories, i.e., is there referential invariance with theory change? Or are there at least three different items, the-Bohr-Rutherford-electron, the mature-Bohr-electron and our-current-electron, i.e., is there referential variance with theory change? Putnam puts the issue in the following general terms:

> What if *all* the theoretical entities postulated by one generation (molecules, genes, etc., as well as electrons) invariably 'don't exist' from the standpoint of later science?'.… One reason this is a serious worry is that eventually the following meta-induction becomes overwhelmingly compelling: *just as no term used in science of more than fifty*

> (or whatever) *years ago referred, so it will turn out that no term used now* (except maybe observation terms, if there are such) *refers*. (Putnam, 1978, pp. 24–25; emphasis in original)

Whether Putnam draws our attention to the meta-induction as a cautionary story to be avoided or a consequence to be welcomed need not detain us; but he does subsequently emphasise that it would be a desideratum of any theory of reference that the argument to such massive reference failure be blocked. Whatever its provenance, PMI has come to have a philosophical life of its own.[1] To set out the induction more clearly, consider the scientific theories θ (relevant to some domain) that have been seriously proposed[2] by scientists over a given period of time. The induction has the following premise:

> For any scientific theory θ seriously proposed at any time t in the past, and whose distinctive theoretical terms were alleged at t to denote some entity, there is some reasonably short span of time N (e.g., 50 years) such that by the end of t + N the theoretical terms of θ were discovered not to denote. (Semantically descending we can say the items in θ's ontology do not exist.)

From this premise we can make the inductive prediction:

> The terms of the theories we currently hold at t = now, will at t + N be shown not to have had a denotation.

We can also make the following inductive generalisation:

> For all theories proposed at any future time t, by later time t + N their theoretical terms will be shown to lack denotation.

From either conclusion we can inductively infer the following pessimistic conclusion: the theoretical terms of our current scientific theories do not denote. Semantically descending we can say that the items postulated in the ontologies of our various theories do not exist. Such a pessimistic conclusion is to be distinguished from a more general kind of philosophical scepticism. Though the conclusion of PMI is close to that of a philosophically based scepticism concerning whether our theories are ever about anything, the considerations invoked based in the history of science are not the usual sort found in premises for arguments about philosophical scepticism. So PMI has a special historical-scientific character.

The conclusions come into direct conflict with a standard conception of an ontological, or metaphysical, realism about scientific entities such as that proposed by Devitt (1997,

[1] There is another version of PMI to be found in Laudan (1981) that differs from the Putnam version in that it puts emphasis on the empirical success of theories not mentioned in Putnam's version of the argument. But the conclusion is much the same in that from empirical success of a theory one cannot reliably infer to any referential success. Laudan's argument has been criticised in Lewis (2001) on the grounds that in arguing from the history of science it commits a sampling fallacy. Neither Laudan's version of the PMI argument nor its alleged shortcomings will be discussed in this paper.

[2] The qualification "seriously proposed" is to prevent the inclusion in the inductive base of frivolously proposed theories that lack minimal epistemic credentials. The induction is intended to be over those earlier theories that (a) featured in the active history of some science, (b) engaged the attention of a number of working experimentalists and theoreticians who made a contribution to their science over some historical period; (c) finally were seriously proposed in that the theories meet minimal epistemic criteria which, if they did not lead to acceptance, at least led scientists to work on them either to show they were false (as in the case of theories of N-rays) or to show that they had some positive evidence in their favour. This qualification is intended to meet an objection, raised by Laudan to an earlier version of this paper, concerning lack of mention of epistemic warrant for the theories in the inductive base.

Sects. 2.3 and 2.4). Ontological scientific realism is to be understood as an overarching empirical hypothesis which says that *most* of the unobservable posits of our current scientific theories exist in a mind-independent manner. The qualification "most", though imprecise, is meant to highlight the point that only foolish realists would claim that whenever our theories postulate unobservables they are *always* right about their existence; realists need to be cautious since our theories are sometimes wrong about what exists. A modicum of fallibilism is an appropriate accompaniment to the formulation of scientific realism.[3] However for the advocate of PMI such qualifications are unavailing since they wish to make the very strong claim that the empirical thesis of realism is refuted; our theories never put us in touch with unobservable existents. Since realism and the conclusion of PMI are contradictory the (fallibilist) realist needs to disarm PMI.

The inductive inference to both conclusions appears to be strong (but as will be seen it fails the *Requirement of Total Evidence*). Assuming this, the challenge to the argument shifts to the truth of the PMI premise. What truth it contains arises from a number of historical cases in which the central theoretical terms of theories have been shown not to denote, for example theories about non-existent items such as celestial spheres, epicycles and deferents, impetus, phlogiston, caloric, electromagnetic aether, and N-rays. A cautiously formulated realism should not be overthrown by these well-established examples of non-existents which have been postulated within science. The qualifier "most", allows the cautious realist to admit these past failures without abandoning ontological realism. The advocate of PMI will reject this defence claiming that realism is still refuted; the number of cases of theories that have postulated non-existent entities simply overwhelms the cautious "most" rendering it ineffectual as a way of saving realism. But can this response be sustained?

Following a point made by Devitt[4] let us suppose that our very early scientific theories concerning some domain (e.g., theories of combustion or light), proposed, say, *c*.1800, were seriously wrong about what unobservables exist (suppose few or none of their theoretical terms successfully denoted). Can we reliably infer that our theories about the same domain are now equally wrong about what exists? Carnap's *Requirement of Total Evidence* bids us take into account all the relevant information before determining the strength of the PMI inference. There is some important missing information concerning methodology that becomes highly relevant. There is an assumption that our current scientific methods for getting us in touch with what exists are no better than they were *c*.1800 for doing this. We ignore the fact that there may well have been methodological improvement and that our methods are now much more reliable for putting us in touch with what exists then they were *c*.1800. So given the relative methodological poverty of some science *c*.1800, the theories proposed in the light of these methodologies were also poverty-stricken concerning the extent to which they put us in touch with unobservables. Let us suppose that by 1900 there was considerable methodological improvement with corresponding improvement of the extent to which scientific theories *c*.1900 were able to put us in touch with unobservables. And

[3] For further qualifications along these lines see Devitt (1997, Chap. 2) and Devitt (2005, Sect. 2).

[4] Considerations concerning methodological improvement are introduced in Devitt (1997, Sect. 9.4) and at the end of Devitt (2005, Sect. 4.2); this important point is employed here.

now, at the beginning of the twenty-first century our methods are quite reliable and our current science does put us in touch with unobservables (*modulo* the fallibilism of the cautious realist). This exposes a weakness in PMI; it ignores *The Requirement of Total Evidence* by remaining silent about a relevant matter, viz., whether or not there has been methodological improvement in various sciences. Granted this omission, PMI looks less promising.

Is the PMI premise true? If one were to take a proper random, statistical sample of theories over, say, a N-year period (e.g., take N = 50 years), it would appear, on a cursory investigation, that the frequency of cases in which there was no loss of denotation at the end of the N-year period would be far greater than those in which there was denotational loss. To illustrate, consider the list of chemical elements developed over the last 200 years or so. Apart from a few classic cases, such as that of the term "phlogiston" at the beginning of modern analytic chemistry (*c*.1800), there have been a very large number of cases in which once a term has been introduced to denote a chemical element it has continued to have successful denotation until our own time. The same can be said of the compounds discovered and named in analytic chemistry; there is hardly a case of denotational failure. Semantically descending, we can say that, within the chemistry of the last 200 years, an element or compound postulated to exist at an earlier time is still an item postulated in our current theories. Much the same can be said of the large number of subatomic particles discovered within physics from the late 1800s; apart from a few well-known examples such as N-rays, there has been much ontological continuity. And a similar point can be made about the kinds of entities postulated in microbiology (bacteria, viruses) and biochemistry. Proceeding differently, it would be possible to vary the length of the period, rather than adopt a fixed N (=50) year period, and take a proper random sample from different lengths of time period, from a few years to a century or more. Sampling over varying time periods hardly alters the verdict just given on the premise of PMI. Such a view is consonant with the idea that alongside our developing sciences there has been improvement in the reliability of the scientific methods used to show that some entity exists.

The premise of PMI is a false generalisation. Converted to a statistical claim the frequency of denotational loss would be low. Combining the failure to observe *The Requirement of Total Evidence* with the suggested counter-evidence from proper sampling across a range of sciences, PMI gives only very weak support to its conclusions, either as an inductive prediction or a generalisation about all theories.

In contrast to PMI, a rival *optimistic meta-induction*, OMI, can be expressed as follows (where the frequency mentioned is to be determined empirically, with the accompanying conjecture that the frequency will be high, or very high):

> For any scientific theory θ seriously proposed at any time t in the past, and whose distinctive theoretical terms were alleged at t to denote some entity, then for any span of time N (e.g., 50 years) the theoretical terms of θ are found, with high frequency, to have the same denotation at t + N as they had at t. (Semantically descending, there is a high frequency of continuity in the items of θ's ontology with any change in theory over time N.)

On the basis of this, one can make a strong inductive prediction concerning the next case of our current theories:

> [W]ith high frequency the terms of the scientific theories we currently hold at t = now, will at t + N be shown to still denote the same entities.

For OMI the strongly supported inductive generalisation would be:

> [W]ith a high frequency, the theoretical terms of theories at any time t, will at a later time at t + N still have their old denotata.

Assuming the historical work does establish the supposed frequencies, the conjecture is that OMI succeeds far more often over periods in the history of science than its rival PMI. This paper considers just one historical case; it will show that OMI is supported in the case of the electron while PMI is not supported, despite the growth in theories about the electron from the early classical theories to the very different Quantum theories.

Both PMI and OMI involve semantic ascent in that they talk of continuity, or lack of it, in the denotation of theoretical terms of some scientific theory. If one is not thoroughly deflationist, so that by means of the schema "'N' denotes iff N exists", one deflates away all questions about denotation in favour of questions about existence, then there remains a substantial matter about the relationship between any theoretical term "N" used in science and items out there in the world. This is a relation that needs to be set out in terms of a theory of denotation. However both PMI and OMI are silent about what that theory might be. In fact, as Putnam points out while introducing the PMI argument, there is, lurking in the background, a semantic theory that gives PMI much of its punch. This is a strongly descriptivist theory of the meaning of scientific terms advocated by Kuhn and Feyerabend and many others. On their account there is a massive amount of incommensurability between theories that provides grist to the mill of PMI; the contextualist, descriptivism of the theory of meaning they adopt entails a rapid turnover of denotata for the same term occurring in only slightly different theoretical contexts. But there is also a contrasting theory of denotation to be found in Kripke (1980), and developed by others, in which descriptivism is downplayed; this theory exploits a much more direct causal connection between, on the one hand, the way a term is introduced and transmitted and, on the other hand, the item denoted. On a more narrow causal approach OMI tends to be supported rather than PMI. Thus it appears that PMI, and perhaps OMI, are not innocent of background semantic assumptions about how the denotation of theoretical terms is to be determined. The upshot of this is that PMI cannot be taken at its face value; it makes concealed presuppositions about how the denotation of terms is to be fixed that, if rejected, undermine any challenge it makes to scientific realism.

Which theory of denotation ought PMI, and OMI, employ? The conditions under which one is to employ a broad contextualist, descriptive approach, or employ a more narrow causal approach, are unclear; for some they are so unclear that they advocate a "flight from reference"[5] and eschew any attempt to invoke reference (denotation) at all. But it is not clear that in doing so PMI is also abandoned. The approach adopted in this paper is descriptivist in character; but it rejects a wholesale descriptivism since this can lead to serious trouble. In claiming this it is not assumed that a theory of denotation can be used to solve problems about existence. This is a matter left to science to

[5] The strategy of a "flight to reference" to resolve issues about what exists is criticised in Bishop and Stich (1998).

determine. But what is claimed is that a suitably crafted descriptivist theory of denotation goes hand in hand with both experimental discoveries in science and theorising about what has been experimentally discovered. To this we now turn.

2. RUSSELLIAN DESCRIPTIONS, RAMSEY SENTENCES AND LEWIS ON FIXING DENOTATION

2.1 Russellian descriptions and a principle of acquaintance

In the following sections it will be shown how Bertrand Russell's theory of definite descriptions can be used to pick out unobservable items involved in experimental and other situations. The task of this section is to set out aspects of Russell's theory and to show that the theory is a special case of Ramsey's account of theories, or more strictly David Lewis' modification of Ramsey's original proposal.

On the classical Russellian theory, a necessary condition for the truth of a proposition, such as [(¶x)Dx]Ax, containing a definite description, (¶x)Dx, is that there is some unique individual denoted by the description; otherwise if no, or more than one, individual is denoted by (¶x)Dx, then the proposition is false. Russell's theory can be generalised to apply not only to individuals but also to kinds. The variable "x" is commonly understood to range over individual objects; but its range can be extended to cover kinds. In this case the description (¶x)Dx denotes the one and only kind K such that it uniquely satisfies the open sentence "D(-)". What a kind may be is something that will be left undefined; all that is assumed is an intuitive grasp of the notion of a kind as illustrated by physical kinds like electrons, chemical kinds like carbon dioxide, biological kinds such as tigers, etc. Finally if a description (¶x)Dx denotes a unique individual or kind, then a name "N" can be introduced for the individual or kind as follows: "N" denotes (¶x)Dx. Such name introduction will be illustrated in the next section for the supposed kind name "cathode ray". In such cases it is the description which carries the burden of fixing a denotation; the name merely serves as a convenient label to attach to the item the description denotes.

Russell put his theory of descriptions to several important uses one of which was epistemological in character. He was (at most times) a realist who thought that we could have knowledge not only of the items with which we are acquainted, but also items with which we are not, or could not be, acquainted.[6] Though his theory of descriptions originated in his 1905 paper "On Denoting" in connection with semantic issues, epistemological issues are not absent. Thus, using one of Russell's examples (Russell, 1956, p. 41), consider the description "the centre of mass of the Solar System at time t". This is a point about which we can say a large number of things in mechanics. But it is not something with which we can be acquainted (say, perceptually), either

[6] In some cases Russell thought that we could know only their extrinsic, structural properties and not their intrinsic properties; this is a matter not discussed here. But see Demopoulos and Friedman (1985) and Demopoulos (2003).

due to our position in the Solar System, or due to the theoretical and unobservable character of such a point. As Russell puts it, we cannot have knowledge of this point by acquaintance, but we can have knowledge of this point by description. Within Russell's epistemology the notion of acquaintance can carry strongly phenomenalistic overtones as when he claims that we are directly acquainted with, for example, the sense experiences to which tomatoes give rise, but we are not acquainted with the tomatoes themselves. However we are rescued from such a strongly empiricist, even Berkeleyan, account of the world; we can come to have knowledge by description about the tomatoes themselves if we cannot get such knowledge by acquaintance. Russell's overall position gives empiricism an important role, but it is not confined to empiricism. He shows how, using his theory of descriptions, we can transcend the items with which we are acquainted, such as experiential items, and adopt a robust realism about external objects of the physical world with which we are (allegedly) not acquainted.

This position is developed in his 1912 *The Problems of Philosophy* when he announces "the fundamental principle in the analysis of propositions": "*Every Proposition which we can understand must be composed wholly of constituents with which we are acquainted*" (Russell, 1959, p. 32, italics in original). For our purposes we need not go into Russell's account of what he means by "understanding a proposition".[7] And we can also set aside Russell's account of acquaintance in which we are only acquainted with sense-data (or universals and possibly ourselves, but not ordinary objects). On a more relaxed position that Russell also adopts in his 1905 paper, we can say that we are acquainted with ordinary objects such as tomatoes. The important step is the manner in which Russell's "fundamental principle" enables us, particularly in the context of science, to move from knowledge of those items with which we are acquainted (suppose these to be ordinary observable objects and happenings in experimental set-ups) to knowledge by description of that with which we are not acquainted (say, electrons, or centres of mass). The important step is made from (a) the items with which we are acquainted and for which we have names and predicates in some language which denote these items, to (b) items with which we are not acquainted but nonetheless we also have names and predicates in the language which denote the items with which we are not acquainted. This step, which takes us well beyond any empiricism embodied in (a) alone, can be made only if we also have at our disposal the resources of logic involving a theory of quantification, variables, logical connectives and the like. Using just these bare, logical resources, and the non-logical terms which refer to, or are about, items with which we are acquainted, we can form descriptions that pick out items with which we are not acquainted.

[7] For more details on this see Demopoulos (2003) who discusses Russell's account of understanding and the constituents of propositions; this is not a matter of significance here. But the use of Russellian descriptions made here is indebted to the story Demopoulos outlines from Russell to Ramsey, Carnap and others. See also Maxwell (1970) who early on recognised the connection between Russell's and Ramsey's theories. In this paper a link is also made to work on theoretical terms by David Lewis, 1983, Chap. 6.

2.2 Russellian descriptions as a special case of the Lewis version of the Ramsey sentence

In presenting Russell's account of definite descriptions in this way, a link can be made to the Ramsey sentence, and more particularly, to Lewis' development of it. Suppose we have a set of statements of a theory T which when conjoined can be expressed as (1), where the bold "**T**" is some function of theoretical terms, t_1, t_2, …, t_n, and observational terms, O_1, O_2, …, O_m, with the whole being equivalent to theory T:

$$\mathbf{T}(t_1, t_2, \ldots, t_n, O_1, O_2, \ldots, O_m). \tag{1}$$

The n theoretical terms commonly appear as kind names, predicates or functors. But as Lewis argues (Lewis, 1983, p. 80) they can be rendered as names of properties, relations, functions and the like, thereby enabling the use of only first order logic in what follows. The m "observational" expressions refer to, or are about, observables (which will include the items with which we are acquainted).

If all the theoretical terms are replaced by variables then one obtains the following open sentence (or Russell–Frege propositional function) containing only variables, logical expressions and non-theoretical or descriptive expressions O_1, O_2, …, O_m:

$$\mathbf{T}(x_1,x_2,\ldots,x_n,O_1,O_2,\ldots,O_m). \tag{2}$$

Denote the Ramsey sentence of (2) by '$\mathbf{T}^R$'. The Ramsey sentence is obtained by placing an existential quantifier in front of the open sentence for each of the variables and forming a closed sentence:

$$\mathbf{T}^R = (\exists x_1)(\exists x_2)\ldots(\exists x_n)[\mathbf{T}(x_1,x_2,\ldots,x_n,O_1,O_2,\ldots,O_m)]. \tag{3}$$

David Lewis' innovation (Lewis, 1983, especially Sects. IV and V) is to show how the n-tuple of objects which uniquely realise the open sentence (2) can have names introduced for them, one at a time, via the construction of a definite description. Thus for the first member of the n-tuple a name "t_1" can be introduced via the generalised definite description on the right hand side:

$$\begin{aligned} t_1 = (\P y_1)[(\exists y_2) \ldots (\exists y_n)(\forall x_1) \ldots (\forall x_n)\\ \{\mathbf{T}(x_1, \ldots, x_n, O_1, \ldots, O_m) \equiv (y_1 = x_1) \& \ldots \& (y_n = x_n)\}. \end{aligned} \tag{4}$$

As set out (4) expresses only the first of n − 1 other descriptions that enable name-introductions for each of "t_2", "t_3", … "t_n". The other n−1 descriptions are easily constructed and can be taken to be part of (4).

Clearly Lewis' procedure differs from that of Ramsey. Moreover it generalises Russellian descriptions in several ways. One way is that it shows how to introduce theoretical terms not just one at a time, but in pairs for two theoretical terms, in triples for three theoretical terms and so on for families of n theoretical terms of some theory T. As such they take into account not only the logical connections between theoretical terms and observational terms, but also the logical connections between the theoretical terms themselves.

For the sake of convenience and clarity, let us prescind from the generality of the above, and consider T to have just one name of a theoretical item, and thus one variable in its open sentence. Also conjoin all the observational terms $O_1, O_2, \ldots, O_m$ and abbreviate the conjunction by "O". Then we have respectively:

$$\mathbf{T}(t,O). \tag{1*}$$

$$\mathbf{T}(x, O). \tag{2*}$$

$$\mathbf{T}^R = (\exists x)\mathbf{T}(x, O). \tag{3*}$$

Lewis' modification of Ramsey's approach is, rather than place an existential operator in front of (2^*) to get a closed existential sentence (3^*), to introduce a definite description operator so that a generalised description is produced:

$$(\P x)\mathbf{T}(x, O). \tag{4*}$$

The last expression says that there is a unique item x such that x satisfies the open sentence "$\mathbf{T}(-, O)$". Also it is possible using the description in (4^*) to introduce some name "t" for the item described; this is simply a special case of the more general expressions of (4).

From this it is clear that Lewis' approach yields a Russellian definite description as a special case. For Russell the expressions $O_1, O_2, \ldots, O_m$ that comprise "O" denote items with which we are directly acquainted. In the context of science we can take these items to be observables. However Lewis imposes no epistemological condition on what can count as the O-terms, $O_1, O_2, \ldots, O_m$. These can be observational, but they could also be *old* or *original* terms whose meaning we already have grasped. Importantly they are O-terms in the sense that the meaning they have is obtained *outside* the context of theory T and not within it; in contrast the T-terms get their meaning within the context of theory T. This liberality in our understanding of O-terms is consistent with Russell's idea of acquaintance, broadly construed, and will be adopted in subsequent sections.

2.3. Two useful modifications: imperfect satisfaction and ambiguous name introduction

A common understanding of the open sentence (2^*) is that the item which satisfies it must be a perfect satisfier; if not then (4^*) cannot be used to denote any item. But this condition can be relaxed in various ways in scientific contexts. To illustrate, consider women A, B, C, D, E and F who have met similar deaths in similar circumstances over a period of time. Then we can form the open sentence " – killer of A&B&C&D&E&F".

The investigators might suppose that there is a unique satisfier of the open sentence and even introduce a name "JR" ("Jack the Ripper") for the killer. On Russell's theory JR is not an object with which the investigators are acquainted, but they are acquainted with A, …, F. However suppose that it is shown that woman F died of causes that could not have been a killing; or that one person killed F while another killed all of the other five. Then has no name "JR" been properly introduced, even if there is a unique killer of the remaining five women? This would be so only if perfect satisfaction is required of the definite description. But we could admit the idea of less than perfect satisfaction and allow a name to be successfully introduced for the item which is the best imperfect satisfier, that is, that item which is a member of some set of sufficiently adequate, imperfect satisfiers and which is the best of this set. In the case described above the name "JR" would then be successfully introduced even though what it denotes is not a perfect satisfier of the open sentence but the best imperfect satisfier.

Such cases readily arise in science. Suppose as in (4^*) term "t" is introduced via the description (¶x)**T**(x, O). But then the laws of T are subsequently altered because they are inaccurate in various ways and T is modified to T^*. For example, the mathematical relations between expressions can be changed in various ways; or a new parameter that was previously unnoticed is added to a law to make it more accurate; or a new law is added to T which makes a difference in certain conditions in which T is applied; and so on. Does this mean that in the earlier theory T, "t" denotes nothing since nothing perfectly satisfies the open sentence **T**(x, O)? To assume so is to go down the path to PMI; but this can be avoided. Perhaps there is some item K which is the best imperfect satisfier of **T**(x, O) in the sense that there is a non-empty set of minimally good satisfiers of **T**(x, O) and that K is the best of these. (K in fact might be the only member of the set. Moreover, if two items K and K^* are tied for first place as best imperfect satisfiers then, as discussed next, any introduced term "t" can ambiguously denote both.) It might turn out that K is a perfect satisfier of the modification $\mathbf{T}^*$(x, O); or K may still be an imperfect satisfier of $\mathbf{T}^*$(x, O), but K satisfies this better than **T**(x, O). The second case opens the possibility that each theory of an historical sequence of theories, T, T^*, T^{**}, etc. can be about the very same item K, where a later theory is a correction of the immediately preceding theory. In such cases ontological continuity under conditions of imperfect satisfaction seems more plausible than failure of denotation throughout the sequence except for the last member when ontological flowering of the real finally takes place (or no such flowering takes place if one accepts PMI). In the past we simply got some things wrong about correctly identified entities. The theory of denotation can be usefully modified to account for such plausible cases of ontological continuity. A term can be introduced by means of an associated open sentence **T**(x, O) (and prior to discovered modifications that give rise to T^*); and its denotation is either the perfect satisfier, or if there is none then it is the best imperfect satisfier, of **T**(x, O). What is important here is that it is the world and its constituent objects, properties, events and the like, which are either the perfect satisfiers of our theories, or their best imperfect satisfiers (under the intuitive constraints envisaged in the modification of T to T^*); or our theories have no satisfiers. Examples of imperfect satisfaction will be encountered in subsequent sections.[8]

[8] For more on imperfect satisfaction see Lewis (1983, Sect. VII) and Lewis (1999, p. 59) where the example used above of term introduction for "Jack the Ripper" is also discussed.

Another modification of the classical theory of descriptions involves the introduction of names which have ambiguous denotation. On the classical theory a description such as "the wife of Ludwig Wittgenstein" does not denote and so cannot be used to successfully introduce a name. In contrast the description "the wife of Frank Ramsey" does uniquely denote and a name can be successfully introduced. But what of the description "the wife of Bertrand Russell"? It does not pick out a unique individual; there were four wives. But does it follow that if some name, say "Xena" is introduced via the description that it fails to denote? The modification proposed is that "Xena" is not non-denoting (as the classical theory would have it) but that it ambiguously denotes each of four distinct women.

There are many cases of term introduction in science in which it is later discovered that the term ambiguously denotes. Such is the case for isotopes in chemistry. Suppose the term "hydrogen" is introduced into science (in whatever way). Once it is discovered that there are several isotopes of hydrogen does it follow that the term "hydrogen" fails to denote? If this is so, then it can be inferred, via the disquotational schema " 'hydrogen' denotes iff hydrogen exists" that hydrogen does not exist. A more plausible alternative would be to claim that there is some denotational contact made with the world when we use the term "hydrogen", but perhaps at a higher level of a genus. If the term "hydrogen" denotes a higher-level genus of element then denotational refinement occurs when the different kinds of hydrogen are distinguished but still using the same name. If there is a need to have different names for these different kinds, the isotopes, then names can be introduced such as the symbols "^{1}H", "^{2}H" and "^{3}H". Unlike other elements, in the case of hydrogen there is a need to have handy proper names for each isotope; so the names "protium", "deuterium" and "tritium" were introduced via denotational refinement. (There are in fact several more isotopes of hydrogen that do not have specific names.)

In subsequent sections a case of denotational refinement will be mentioned, but such refinement plays no role in the case of the terms "electron" or "cathode ray". It would appear that when these terms were introduced a fundamental kind of thing was named, so there has been no need for denotational refinement. But the world might, one day, tell us that denotational refinement is in order and that there are, currently unknown to us, different kinds of electron with different internal structures and thus different properties. To conclude that we had not been talking about anything when we used the term "electron" is to be misled by too limited a view of how denotation is fixed. Rather, we had made some appropriate denotational contact with the world; but in the light of the discovery of electrons with different internal structures, our term "electron" was, unbeknownst to us, ambiguous (or had no determinate denotation) and that denotational refinement would be in order.[9] We now turn to an application of the above semantic theory to an episode in physics.

9 See Lewis (1999, p. 59) who also advocates the idea of what he calls "indeterminacy of reference" following earlier work of Hartry Field (1973) on this. For a further scientific example, a similar story about Lorentz's use of the term "ion", as Theo Arabatzis points out to me (private correspondence), underwent referential refinement. Initially it would have ambiguously referred to the ions of electrolysis but later underwent refinement when, after Zeeman's discovery, he realised that there were important differences, so important that *c.*1899 he started to call them "electrons"; see Arabatzis (2006, pp. 84–85).

3. JULIUS PLÜCKER'S EXPERIMENTS WITH GEISSLER TUBES

Experiments on the passage of an electric current between electrodes in a closed tube had begun in the first decade of the 1700s and for the next 120 years a well-recorded sequences of phenomena were noted as the pressure of the gas inside was reduced. In the late 1830s Faraday pushed the limits of the then available electricity sources and vacuum pumps reaching a minimum pressure of approximately 1 mm Mercury. At about this pressure for tubes filled with air, a violet glow appears at the cathode which pushes the pink glow already in the tube towards the anode; the violet and pink glows do not touch and are separated by a dark space. This Faraday investigated and is now known as the "Faraday dark space". These limits of the then available experimental apparatus were transcended when the Rühmkorff induction coil was devised in the early 1850s to produce high-voltage electric currents, and when Geissler invented in 1855 a quite new kind of mercury pump that would enable experimenters to reach entirely new levels of low pressure in tubes. In 1857 he also invented a new kind of discharge tube called by Plücker the "Geissler tube"; these were of various shapes with electrodes fused into them and filled with various kinds of gases. The coil, pump and tubes were a technological breakthrough in the investigation of electric discharge through gases and quickly became standard equipment in physics laboratories. Any experimenter using them could reliably produce the effects already observed and then proceed to make novel investigations. Julius Plücker was the first to do just this.

Some of the phenomena to be observed at the newly available low pressures are as follows. The pink glow on the side towards the anode, breaks up into a series of striations with the concave surfaces facing the anode. At even lower pressures the violet glow around the cathode breaks into two parts with a new dark space emerging between them, now known as the "Crookes dark space" (owing to later investigations by William Crookes). At lower pressures the Crookes dark space grows in size pushing the Faraday dark space and the striations towards to anode until the Crookes dark space fills the tube and there is no luminosity. At about 0.01 mm Mercury (about 1/100,000th of an atmosphere) a completely new phenomenon appears: a greenish-yellow glow bathes the walls of the tube. This usually appears on parts of the tube away from the cathode; but in the case of Plücker's original experiments these were close to the cathode owing to the peculiarity of the tube he used.[10] It was this new phenomenon that excited the interested of many experimentalists, Plücker being the first to record them in papers of 1857–1859 of which Plücker (1858) is an English translation of one paper.

This new phenomenon is a reliably reproducible effect that many could bring about in their laboratories. The production of this, and other phenomena Plücker observed,

[10] Dahl makes the following interesting comment on Plücker's experiment: "Apparently the fluorescence, 'the beautiful green light, whose appearance is so enigmatic' was not uniformly spread over the wall of the tube, as is usually the case in a modern discharge tube when the Crooke's dark space attains a maximum. Instead, probably due to some quirk in tube construction in Plücker's experiments, it was concentrated in patches near the cathode. But for this fortuitous quirk, Plücker would not have discovered that the position and shape of the fluorescent patches are altered by a magnetic field" (Dahl, 1997, p. 54).

can be set out in the form of an experimental law E based on the disposition, under conditions C, of the experimental set up to produce the range of effects P_i (labelled "P" for Plücker who observed or experimentally produced the effects).

> (E) There is a repeatable experimental set-up concerning the apparatus (viz., Plücker's Geissler tube) which, under conditions C of its electrical source and sufficiently low pressure beyond the emergence of the Crookes dark space, has a law-like disposition to produce phenomenon P_i.

3.1. Plücker's observations

What are the effects that Plücker observed in conditions C? The first is simply the new phenomenon itself which he refers to as a "beautiful green light, whose appearance is so enigmatical" (Plücker, 1858, Sect. 35, p. 130):

> (P_1) There is, in conditions C, a coloured glow (in Plücker's experiment a greenish light)[11] on the glass of the tube (usually at the end opposite the cathode).

Plücker was aware of Davy's experiment in the 1820s which showed that the shape of an electric arc (produced by separating two electrodes initially in contact) could be affected by a magnet. Since he believed there was something akin to a stream of electric current in the tube, then it should also be similarly deflected. So he conducted a number of experiments by placing different kinds of tubes with different gases in different orientations with respect to a magnet. Plücker gives lengthy qualitative descriptions of how a magnetic field affects the light in the tube before the Crookes dark space appears. More important for our purpose is what happens to the "enigmatical beautiful green light" which appears on the inside contours of the glass after the Crookes dark space fills the tube. It can be moved back and forth as the polarity of the surrounding magnet is changed. This is Plücker's central experimentally manipulated effect:

> (P_2) In C, the patches of coloured glow can be moved by a magnetic field.

When there was an electric discharge through a Geissler tube, Plücker believed that what he called "rays of magnetic light" radiated from every point of the surface of the cathode. So he coated all of a cathode except a point-like extremity with glass and noted the single stream of light that emanated from the point. Owing to its concentration at a point the light can become visible (ibid., Sect. 30, p. 131). On the basis of this he drew an analogy with what would happen to iron filings placed near a point-like magnetic field; the iron filings would all move into a single line along the line of magnetic force emanating from the point. Similarly for the "magnetic light"; all the rays of light passing through a point would be aligned. But as he makes clear, this is analogy only between the iron filings and the "rays of magnetic light" and not an account of

[11] The colour of the glow depends on the chemical nature of the glass, in this case the yellowish-green colour being due to soda glass; lead glass gives blue, etc. This is something others, such as Crookes, discovered later; the particularities of colour play no role in the story being told here.

the nature of the "magnetic light" itself (ibid., Sects. 47–50, pp. 408–409).[12] This we may sum up as follows:

> (P_3) A point-like cathode produced a visible beam and a less diffuse more concentrated glow effect.

Also the following manipulation effect can be observed as a special case of P_2:

> (P_4) With a point-like cathode, the ray of light, and the patch of coloured glow it causes on the tube, can be deflected by magnet.

Plücker also reports that the coloured glow inside the tube does not depend on the position of the anode; and all the effects he observed were independent of the metal used as electrodes (usually platinum but often coated with other metals):

> (P_5) The glow is independent of the anode position.
> (P_6) The glow is independent of the metal used as cathode and anode.

What we now know to occur in Geissler tubes is well described by Weinberg:

> We know now that these rays are streams of electrons. They are projected from the cathode by electrical repulsion, coast through the nearly empty space within the tube, strike the glass, depositing energy in its atoms which is then readmitted as visible light, and finally are drawn to the anode, via which they return to the source of electricity. But this was far from obvious to nineteenth century physicists. (Weinberg, 1990, pp. 22–23)

The stream of negative electrons from a finely pointed cathode produce a ray of light; this is so because, even at low pressures of the contained gas, the concentrated stream of electrons still manages to hit the gas molecules thereby emitting light. The stream of negative electrons also repel one another as they pass down the tube; hence the "rays" from the fine point of the cathode are splayed out to a small extent yielding patches of coloured glow inside the glass of the tube. Later experimenters were able to place a screen inside the tube that would detect the electrons that grazed it as they passed down the tube, thereby showing more clearly the beam and its deflection by a magnetic field. However for the remainder of this paper we will eschew the whiggism of considering the experimentation and theorising from our current point of view and remain in the context of the late nineteenth-century physics when these matters were not obvious.

3.2. Constructing an identifying description denoting the cause of what Plücker observed

What philosophical issues arise out of this episode in the history of science? One of the main claims of this paper is that Plücker's investigations provided sufficient information to yield identifying conditions for an entity of some kind which caused

[12] As Plücker says: "The rays proceeding from this point collect in one single line of light, which coincides with the magnetic curve passing through the end of the negative electrode, and which luminosity render such magnetic curve visible. Thus every ray which is bent in this magnetic curve, forming a portion of the arc of light, behaves exactly as if it consisted of little magnetic elements placed with their attracting poles in contact.… By the above illustrations I have merely sought to make the nature of the phenomenon intelligible, without in the least attempting to describe the nature of the magnetic light itself" (Plücker, 1858, Sects. 49–50, p. 409). By speaking of "as if", Plücker is not proposing a theory of what the "rays of light" are, viz., a thread of magnetic elements or anything like that.

what Plücker observed, even though no intrinsic property of the kind was known of the entity but only its extrinsic properties in specified experimental set-ups.[13] Moreover this identification, and re-identification, can occur in the absence of any name introduced to refer to the kind, as was the case for Plücker. Consider the kind presupposition first.

On the basis of the Causal Principle, we can say that the observed phenomena P_1 to P_6 are caused by an otherwise unknown "something" or "somewhat" of a particular kind; call this kind "K" (whatever it be). What is clear is that instances x of K arise in each experimental set-up causally bringing about P_1 to P_6. This involves a kind presupposition of the following sort:

> (K) There is a kind K (unique or not), instances x of which arise in the conditions of E, and which are casually efficacious in bringing about a range of phenomena such as P_1 to P_6.

We can leave open what kind does the causing. The "somethings" were variously regarded as flow of electricity (whatever that may be), or beams of light (though talk of light can be misleading) or rays, though the connotations of these terms are not, at this stage, important. We can also leave open what ontological category K belongs to, e.g., a kind of substantial object such as particles, corpuscles or whatever; or events or processes such as electromagnetic waves in an aether; or any other ontological category one might like to consider. Such a bare causal assumption is unproblematic; clearly there must be something giving rise to the effects, and not nothing, unless the world is much more indeterministic than we think. Fairly direct evidence that there is a cause at work bringing about the phenomena is shown when the current from the Rühmkorff coil is turned off and on.

But there need not be just one kind of thing that does the causing. It could be that there is, say, a single genus each species of which is, or can be, causally involved. Here the kind presupposition is not dropped but moves to a higher taxonomic level. In the next section an example of a term introduction, that of "canal ray", will be given in which the supposition that there is a unique kind named is dropped in favour of a genus without claiming that there are no canal rays. Here the notion of ambiguous name introduction and indeterminacy of denotation introduced in Sect. 2.3 comes into its own. More extremely, presupposition (K) might be challenged in cases where the causes turn out to be quite heterogeneous and there are several kinds giving rise to some effect. As an example consider the atmosphere which we identify by the casual role it plays in giving us certain feelings on a windy day, sustains our lives when we breath it, sustains fires, and the like. Though it was initially thought to be a single kind of substance, we now know it to be a heterogeneous collection of different kinds. But in discovering this we did not conclude that there is no atmosphere. Here the idea of denotational refinement of Sect. 2.3 once more plays an important role concerning the

[13] It is important to note that, in one sense, we have always been in casual contact with electrons, as when we are struck by lightning, or when earlier investigators experimented on electricity. But none of this was used as away of providing information about identifying a "something" and manipulating it for various purposes. It is Plücker's manipulation-based identification that puts us into contact with electrons in ways that are other than bare casual contact.

terms we use to talk about the atmosphere and talk about the heterogeneous kinds that make it up.

In what follows it will be assumed that (K) can be understood in a sufficiently broad way to include cases where there is no one kind involved and not be undermined by the involvement of multiple kinds. In the case of Plücker's discoveries we now take him to have been performing experiments on streams of electrons. So far as we know, there is just one kind of electron. So in this case supposition (K) has been a correct one to make. But even if it turns out that there are different kinds of electrons, or they have a heterogeneous character and (K) has to be modified in some way, it would not follow that there were no electrons; rather they are quite different from what we took them to be.

Suppose that we are given (E), (K) and Plücker's observable and experimental effects P_1 to P_6 (in what follows abbreviate the conjunction of the six P_i effects to "P" for Plücker). Then we can reconstruct the manner in which Plücker and others were able to employ a definite description to identify a kind involved in the repeatable experimental set-up governed by (E), proceeding as follows.

Form an open sentence of the following sort (where "–" is a blank for a variable ranging over kinds):

(1) – is a kind and instances x of the kind – in condition (E) are such that x cause effects P.

There are three possibilities to consider concerning the satisfaction of the open sentence. First, there is no kind of thing which satisfies the open sentence (i.e., nothing is true of the open sentence). This alternative can be set aside; it would arise when either no kind of thing at all satisfies the open sentence, or if any particular things do satisfy it, they are quite different from one another and do not even form heterogeneous kinds at all, as in the case of the atmosphere. Second, there is exactly one and only one kind of thing that realises it. In what follows we will assume that this is the case. Third, two or more kinds realise the open sentence. In such a case of heterogeneity, we cannot say that nothing satisfies the open sentence (the first alternative); rather there is nothing akin to a single natural *kind* that satisfies it. Later we will return an example where two or more kinds of thing realise such an open sentence.

If we put an existential operator in front of the open sentence then we get a particular instance of a Ramsey Sentence. Thus where "k" ranges over kinds we have:

(2) There exists a (unique) kind k such that instances x of k in conditions (E) cause effects P.

However such a Ramsey sentence does not say that there is one and only one kind that satisfies the open sentence. But if we adopt the stance of David Lewis's modification of the Ramsey sentence suggested in Sect. 2, then we can form a definite description which picks out a kind (where "(¶ –)" is the definite description operator):

(3) (¶ k)[for instances xs of k in experimental set-up (E), the xs cause effects P].

This is a generalised version of a Russellian definite description, but in this case it is not a unique individual object that is being picked out but a unique individual kind.

Given this description, we are now in a position to introduce a name "K" for the kind specified in the definite description above:

(4) Let "K" denote (¶ k)[for instances xs of k in experimental set-up (E), the xs cause effects P].

Plücker in fact introduced no such proper name – this was something done by his followers. However he did use the phrase "rays of magnetic light" to capture the fact that the light emanating from a point-like cathode could be manipulated by a magnet. The above reconstruction shows that there was available to Plücker and his contemporaries an identifying description that does pick out a kind. What the above makes clear is the main burden of denotation fixing falls on the identifying description and not the name which is introduced on the basis of the description.

One important feature of the definite description in (3) and (4) is that identifying reference is made through purely extrinsic, or relational, properties involving the experimental set-up in which instances of K arise. Nothing is indicated in the description about what intrinsic (non-relational) properties instances of K possess. Nor is anything indicated about its nature or what ontological category it might belong to (substance, property, event, etc.). Later experimental work came to tell us about some of the intrinsic properties of kind K; but knowledge of any of these is not necessary for identifying reference to occur, as (3) shows.[14]

It is the world which determines whether or not there is some (unique) kind which satisfies open sentence (1), and thus whether the description formed in (3), or the name "K" introduced in (4), picks out a kind. The item which satisfies the open sentence may be a perfect satisfier of the open sentence. Or it might be an imperfect satisfier which is the best of a sufficiently satisfactory set of satisfiers. Where there are two or more satisfiers, either perfect, or imperfect best, is a matter to be considered further in the light of the idea of ambiguous denotation. In the above example perfect satisfaction seems more likely than imperfect best satisfaction since the extrinsic properties specified in the denotation fixer are broad enough to ensure that something is captured by them – assuming the world is, in this case, generous to us and does contain a unique kind that fits the description, and is not so ungenerous as to contain only a quite heterogeneous jumble of objects which fit the description.

What is the status of the name-introducing claim? Since it is an introduction of a name "K" for a kind, we know a priori that this is true (at least for those who introduce term "K" and their immediate community to whom the use of "K" is transmitted). But it is not a truth of meaning; nor does it specify the intension expressed by "K", or give us the concept of a K. Moreover it is a quite contingent matter that samples of the kind K ever get involved in Plücker's experimental set-ups with its Geissler tubes and the like. In some possible worlds the course of physics might have been quite different

[14] Some might see in this support for versions of structuralism as advocated in Russell, (1927, pp. 226–227 and 249–250), a position criticised in Newman (1928), and reviewed more recently in Demopoulos and Friedman (1985). The view is also found in Maxwell (1970, especially p. 188) where it is argued that from the bare Ramsey Sentence something like structuralism follows. Though it will not be argued here, from the fact that only extrinsic properties are used in denotation fixing, nothing follows about structuralism. In fact it will be argued that much later, especially with the work of J. J. Thomson and others, we did come to know something of the intrinsic features of what has been extrinsically identified, viz., cathode rays or electrons.

and there never had been discharge tubes, or the inhabitants of that world never went in for physics; yet there are exist Ks in that world. A quite different matter to consider is whether, if electrons do behave the way they do in Geissler tubes, then (contrary to Hume) it is necessary that they bring about effects P; that is, it is not metaphysically possible that there be Ks (electrons) in discharge tubes in the experimental set-up (E) but they not produce these effects. This question does not have to be answered here. The main concern is the Kripkean point that the name-introduction claim is known a priori but it is not a truth of meaning; but this does not decide whether it is a metaphysically necessary truth.

Of the effects P_1 to P_6 that comprise P the most central ones are P_1 and P_2, and their corollaries P_3 and P_4; these involve experimental manipulation of the coloured glow. In the above we have supposed that there is one kind of thing that is being manipulated.[15] But effect P_2 is more than a bare causal claim; it tells us that kind K has the disposition to be affected by magnetic fields (but in the absence of any knowledge of how the disposition works and any causal mechanism at work). It is their susceptibility to such manipulation that underlies realism about what is being manipulated. Also the manipulation condition underlies much subsequent quantitative experimental work (sometimes combined with theory) about what is being manipulated.

P_1 and P_2 also need to be taken in conjunction. Suppose there is a "something", a k kind, instances of which bring about effect P_1, and they also have the underlying disposition to be affected by a magnetic field as in P_2. These two features are central in identifying the "something". In contrast, suppose there is a possible world containing k and y kinds such that instances of the y kind in C also cause a glow effect but they are not deflected by a magnetic field; we can conclude that the y kind is not the same as the k kind. Again there is a possible world containing k and z kinds such that instances of the z kind in C are deflected by a magnetic field but they do not cause the glow effect; then the z kind is not the same as the k kind (or the y kind). To the best of our knowledge these possible worlds with their y and z kinds are not our actual world; our actual world contains just the k kinds and they both cause the coloured glow effect and are disposed to manipulation in magnetic fields.

3.3. An important restriction on what is to be admitted into identifying descriptions

Identifying description (3) of the previous section would provide any reader of Plücker's paper with a recipe for constructing similar experimental conditions, observing the same phenomena – and so being in the presence of the same unobservable "something", the kind K, that Plücker encountered. However not all the claims that Plücker makes in his paper are included in the denotation fixer. What have been excluded are

[15] Hacking's remark "if you can spray them, then they are real" (Hacking, 1983, p. 22) underlies a realism about what we use when we manipulate (in this case the magnet and its field). But there is also what is manipulated (in this case the "something" that causes (P_1), etc.), and this too must be as real, though not much may be known about it. Causal relations can only hold between the equally real "what is used to manipulate" and "what is thereby manipulated".

speculative explanatory hypotheses about what is going on in the tube. Consider the following sputtering effect that Plücker noted (Plücker, 1858, Sects. 3 and 51). During the electric discharge he noted that small particles of the platinum metal of the cathode were torn off and deposited on the glass of the tube in the vicinity of the electrode. Tests were made to prove that the deposit was platinum metal from the cathode. He also noted that the glass was not always so blackened; but when it was there was no tendency for the free cathode particles to move towards the anode. He speculates: "It is clearly most natural to imagine that the magnetic light to be formed by the incandescence of these platinum particles as they are torn from the negative electrode" (ibid., Sect. 51, p. 410).

This we may set out in the following two-part condition S (for sputtering), of which (i) is the report of observations, and (ii) is a causal explanatory claim:

> (S) (i) in (E) bits of incandescent platinum are sputtered off from the cathode and deposited on the glass; and (ii) the occurrence of (i) is the cause of the " 'rays of magnetic light".

Clause (ii), unlike (i), introduces a speculative causal explanatory hypothesis. Should this be included in any denotation fixer? It would be by those who understand Ramsification to apply to whole theories. In this case one could take all of Plücker's observational and explanatory claims in conjunction. Then one could consider the following denotation fixer: the unique kind of thing k such that instances of k satisfy both (P&S). If so, then unlike denotation fixer (3), nothing satisfies both (P&S). The reason is that S is false; the occurrence of (i) is not the cause (ii). As will be discussed in Sect. 4.2, Goldstein showed that the sputtering effect is not due to anything given off from the cathode; rather it is due to a flow of the positively charges ions of the residual gas impacting on the cathode surface which gouge out small bits of the metallic electrode. To include false claim S along with P in any denotation fixer would ensure that no denotation is fixed. So in order to fix a denotation, such speculative causal explanatory hypotheses should be omitted from any definite description; it should employ only terms referring to observational properties, or experimentally determined properties. This raises the matter of what should be included in any denotation fixer, and what excluded.

Should one use in denotation fixing descriptions only terms which denote observables? To do so would be too restrictive. One needs to include terms which denote experimentally determined properties, but also allow for theoretical terms as well (e.g., "charge", "mass", etc.). Laws can also feature in generalised denotation fixing descriptions; such laws need not be restricted only to relations between observables, e.g., the law about the constancy of the ratio of mass to charge (m/e) for electrons. Importantly background theories are often used to experimentally determine such law-like relations between properties, which in turn are used in fixing denotations. (An example of this is given in Sect. 6.1, Thesis (B) on the background theory employed to determine the m/e ratio which is then subsequently used to identify electrons.) If these experimentally determined relations are employed in any denotation fixer, they will bring in their wake an involvement with theory. So the common theory/observation distinction is not the right one to invoke when restricting what terms can go into denotation fixing descriptions.

More promising is a distinction between those items which do not feature in explanations and those which do. Physicists often draw a distinction between the phenomenological and the fundamental; it is the phenomenological which stands in need of explanation but does not itself explain, while the fundamental does the explaining of the phenomenological.[16] Alternatively the distinction can be drawn in terms of observables and experimentally determined properties and relations versus theoretical models which purport to explain what we observe or experimentally determine. Though the distinction is not sharp, it is some such distinction which underpins the distinction between those terms which go into denotation fixing descriptions and those which do not. A related distinction is sometimes made between *detection* properties versus *auxiliary* properties. The terms in a denotation fixing description are about the properties used to detect some item (such as the properties used to pick out cathode rays); but other non-detection properties do not perform this function but may have other auxiliary functions such as specifying how the item picked out by detection properties behaves in other respects, or how it explains what we can detect.[17]

The significance of a distinction drawn along the above lines is that it precludes forming a denotation fixing description which incorporates the total context in which a term occurs, as do most accounts of the use of the Ramsey Sentence. In such cases it is much more probable that no denotation is fixed than if a more restricted context is drawn upon. In what follows, the context will be restricted to those which are not explanatory and pertain to what we can observe or experimentally determine. So restricted, the denotation fixing description is less likely to pick out nothing, or to carry excess explanatory baggage which takes us in the direction of incommensurability.[18]

4. HITTORF, GOLDSTEIN, CATHODE RAYS AND CANAL RAYS

4.1. Hittorf's Maltese Cross

For our purposes, it was Johann Hittorf who, following Plücker's experimental recipe, made the next important advance in our knowledge of what goes on in Geissler tubes. He placed a solid body, for example one in the shape of a Maltese Cross, between a point-like cathode and the glow on the walls of the tube. He noted that a shadow of the body, a cross, was cast on the tube's walls. He inferred from this that the glow on the tube is formed by a "something" coming from the point-cathode, travelling in straight lines in a cone with its apex at the point-like cathode (assuming of course that there is no magnetic field present), and impacting on the walls of the tube causing the coloured glow. The paths of these "somethings" can be blocked by a body causing its shadow to appear on the walls of the tube which otherwise exhibit Plücker's enigmatic green glow. We can sum this up in a further clause (H), for Hittorf:

[16] Such a distinction is drawn in Cartwright (1983, pp. 1–2), though here a greater emphasis is put on what does and does not explain.

[17] The distinction between "detection properties" and "auxiliary properties" is made in Chakravartty (1998, p. 394). The distinction is put to much the same purposes as those required in this paper, viz., to enable the fixing of denotations independently of other matters such as theoretical explanations.

[18] Restricted versions of the Ramsey Sentence have been proposed by others; the version in Papineau (1996) is the one most congenial to the purposes of this paper.

> (H) The "somethings", xs of kind K, travel in straight lines from a cathode to the walls of the tube causing its glow; and if the cathode is point-like, and xs are intercepted by an opaque body, then a shadow is cast on the walls of the tube.

What Hittorf shows is that the kind of "something" that Plücker identified, viz. (¶ k)Pk, has a further set of properties – call these "the H-properties". That is, we can now claim: [(¶ k)Pk]Hk.

From our point of view, the other important thing that Hittorf did was to introduce a name of the "somethings" of kind K. We can reconstruct his name introduction in the following way:

Glimmstrahlen (glow rays) = (¶ k)Pk.

This is not a Kripke-style baptismal reference to a kind via pointing to some perceptually available samples in one's vicinity and then introducing a kind name for anything that is of the same kind as the samples. This is not possible since the kind, whatever it is, is not directly observably accessible to us. All that is accessible are the properties we can observe in some experimental set-up. By means of our acquaintance with these properties a generalised Russellian description can then be formed that takes us to the kind with which we are not acquainted.

Note also that this name might convey some connotations that are unwarranted by the descriptive denotation fixer on the right-hand side. That they are "glow" rays is perhaps a connotation due to their causal effect when impacting the glass of the tube. What connotations the term "ray" contains might vary according to one's theory of what rays are; but this need not be obtrusive in this term introduction. Both particle and aether theorists talked of "rays" of light, the term "ray" being neutral between the two theories. However some physicists might wish to build more into the concept of a ray than this, in which case the connotations of "glow rays" goes beyond what the description (¶k)Pk contains.

4.2 Goldstein's two discoveries and name introductions

The next salient advance was made by Eugen Goldstein in 1876. He discovered the following, which we can lump together to get clause (G) for Goldstein. He showed that shadows were cast by an opaque object not only when the cathode was point-like, but also when the cathode formed an extended surface. In the latter case a shadow of an object was cast on the tube only when the object was close to the cathode; if the object was too far from the extended cathode then the shadow edges would become diffuse and not sharply defined, or no shadow would be formed at all. Goldstein also showed that whatever came from the cathode was not emitted in all directions but was largely at right angles to the surface, and travelled in straight lines (in the absence of a magnetic field). In this respect what is emitted from the cathode behaves differently from the light that is emitted from an incandescent lamp. He also definitively established a result that Plücker proposed – whatever is emitted from the cathode is the same regardless of its metal composition. These were important additional discoveries[19] that we can incorporate into clause (G) as follows:

> (G) The "somethings" of kind K are emitted at right angles to the surface of the cathode and travel in straight lines and are not emitted in a spectrum of angles (as in the case of light); and they are independent of the material of the cathode.

The new knowledge we have of kind K is: [(¶ k)Pk]Gk; and combining the knowledge obtained by Hittorf and Goldstein we have: [(¶ k)Pk](Gk&Hk).

In 1876 Goldstein introduced another name to refer to the "somethings" in the following way:

Kathodenstrahlen (cathode rays) = (¶ k)Pk.

This new name has connotations of its own some of which arise from the denotation fixer. Thus "cathode" tells us something about the extrinsic, causal origin of the "somethings", unlike Hittorf's that tells us only about their effect. But "ray" is still vague and may convey connotations that go beyond the denotation fixer according to what theory one adopts of the nature of the rays. What is known of cathode rays remains largely its extrinsic properties in experimental set-up E, and hardly anything intrinsic.

The term "cathode rays" caught on while that of "glow rays" did not. Whichever name we adopt, why do we want a name to refer to kinds of thing? As is evident, to have continued to use any of the above descriptions, such as (¶x)Px, would have been rather cumbersome. In addition, to keep talking about *the "somethings" we know not what* that arise in the experimental set-up is hardly helpful. Our conversation and communication is vastly improved if we can have names for individuals and kinds. And this is true of the community of scientists. Hence their need to coin a name to refer to the kind of thing they were all intent on investigating. Other names were also used in the nineteenth century. For example, John Peter Gassiot (a collaborator with Faraday in the late 1850s whose flagging interest since the late 1830s in discharge tubes had been rekindled by Plücker's work), had coined the term "negative rays" (Shiers, 1974, p. 93). But the name for the kind, and some of its connotations both relevant and irrelevant, is not as important as the description that fixes the denotation for the name.

Goldstein is also responsible for the discovery of Kanalstrahlen, i.e., canal rays (also known as "positive rays"). These play a small role in the story being told here since they concern the sputtering effect that Plücker had observed and which Goldstein investigated. One issue confronting experimenters was the problem of the suspected, but missing, positive counterflow in the direction from anode to cathode. This proved difficult to detect until Goldstein in 1886 devised a way. He bored one or more holes in a solid plate-like cathode through which the positive flow might pass through the plate and then beyond (on the opposite side from the anode), where they can be detected. The holes in the plate were to function like a duct or channel (in German "Kanal") for the flow rather than blocking the flow. In Goldstein's apparatus, at low enough

[19] For an account of this work see Dahl (1997, p. 57).

pressures the characteristic cathode ray effects arose; but on the back side there were long columns of yellow light: "The yellow light consists of regular rays which travel in straight lines. From every opening in the cathode there arises a straight, bright, slightly divergent yellow beam of rays" (cited in Dahl, 1997, p. 81; Magie, 1935, p. 577). Unlike cathode rays, the light is readily visible; and its colour varies with the gas used in the tube. One important difference was that they were unaffected by a magnetic field just strong enough to deflect cathode rays.

These features are sufficient to provide a description of a repeatable experimental set up, E^*, and a set of claims (G^*) (for Goldstein) which can then be used to form a description (¶x)G^*x that supposedly denotes a unique kind of item. Goldstein tentatively introduces a name for the kind on the basis of the following extrinsic description: "Until a suitable name has been found, these rays, which we now cannot distinguish any longer by their colour, which changes from gas to gas, may be known as 'canal rays' ['Kanalstrahlen']" (loc. cit.). Put more formally in terms of the theory of name introduction by means of descriptions we have: "canal rays" denotes (¶x)G^*x. The name was adopted for a considerable time; but its connotations are based on a highly contingent feature of the "canal" rays, viz., the channelled holes in the cathode through which they pass before detection.

Goldstein's discovery remained in limbo until later experimenters, from Wilhelm Wien in 1898 onwards, were able to deflect them in strong magnetic fields. It turned out that different samples of canal rays could be deflected through different degrees. We now know canal rays to be a quite broad genus, positive ions, of which there are many species, each having a different angle of deflection in the same field.[20] We now know that the smallest is a proton which has the greatest degree of deflection, or as Stark calls them "hydrogen canal rays"[21] when hydrogen was the gas in the tube. When helium is the gas in the tube then "helium canal rays" are produced; and so on. Here it would be misleading to conclude that there was no denotation for the name "canal rays" since a large number of different kinds fit the description; rather it turns out that the term Goldstein introduced either picks out a broad genus, or it names ambiguously a number of different kinds. Finally Goldstein's discovery of canal rays showed that Plücker's causal explanatory hypothesis about sputtering was quite wrong.

5. IDENTIFICATION AND RIVAL THEORIES AND MODELS OF CATHODE RAYS

Given the identification of cathode rays and the advancing knowledge of their properties, such as that contained in [(¶ k)Pk](Gk&Hk), which was obtained by observation or experiment, it is now possible to advance theoretical and/or explanatory models of

20 See Dahl (1997, pp. 80–81 and pp. 265–266) on Goldstein's discovery of canal rays. Extracts from Goldstein's 1886 paper appear in Magie (1935), pp. 576–577.

21 See the 1919 Nobel Prize lecture of Johannes Stark (1967, p. 430) who did experimental work on canal rays in the early 1900s and discovered the Doppler effect for canal rays; he also worked on hydrogen canal rays, helium canal rays, etc., including molecule canal rays depending on the molecular gas in the tube.

the behaviour of cathode rays. These models require the prior identification of cathode rays and do not contribute to their identification. Goldstein was an advocate of models in which cathodes rays were some kind of electromagnetic wave propagation in an all-pervasive aether. Others advocated models in which they were some kind of particle. It seems that the first person to claim this was Cromwell Varley in 1871 who said that they are " composed of attenuated particles of matter projected from the negative pole by electricity in all directions, but that the magnet controls their course" (cited in Dahl, 1997, p. 62). Nothing is said about the particles concerning their size or mass; but this is the beginnings of a simple model in which cathode rays are streams of small particles.

A similar view was developed by William Crookes who investigated the dark space now named after him, making a not implausible hypothesis about what was happening in it. He suggested that the molecules of the remaining gas come into contact with the cathode and acquire from it an electric charge. They are immediately repelled from the surface of the cathode at right angles because of the mutual repulsion of like particles. This would produce a stream of moving molecules of the gas at the very low pressures of the gas. Crookes used this hypothesis to explain the growing Crookes' dark space that appears in the tube. He also suggested that one might be able to measure the mean free path of the molecules of the gas at such a low pressure in the dark space, and speculated further on how this theory might explain the brightness that emerged at the end of the Crookes dark space away from the cathode:

> The extra velocity with which the molecules rebounded from the excited negative pole keeps back the more slowly moving molecules which are advancing towards that pole. The conflict occurs at the boundary of the dark space where the luminous margin bears witness to the energy of the collisions. (Cited in Whittaker, 1951, p. 352)

According to Crookes, the dark space is dark because there are no collisions occurring in it; and the bright space at one end of the tube is bright because that is where the collisions take place. But some of the charged molecules do get through and cause the characteristic yellowish-green glow on the glass of the tube; the lower the pressure the more get through until it is at a maximum when the Crookes' dark space fills the tube.[22]

Crookes' theory entails the idea that the charged molecules can exert a pressure, a consequence he developed both theoretically and experimentally. The experimental demonstration consisted of the radiometer fly developed largely by Crookes' laboratory technician Gimingham (see Dahl, 1997, p. 72). It comprises a horizontal cross suspended on a steel point so that it could rotate freely. Attached to each arm of the cross was a small thin mica disc. If one side of the disc was blackened and the other left shiny then when it was exposed to bright light the cross rotated suggesting that there was pressure due to the light exerted on the mica discs. Crookes got Gimingham to adapt this for one of his "Crookes' tubes" (his version of Geissler tube) in an experiment to show that his radiometer would rotate. When placed on some glass rods aligned like railway lines the fly easily roll along them. When the charges were

[22] For an account of Crookes' ideas see Whittaker (1951, p. 352).

reversed on the anode and cathode the fly could be made to roll back the other way (see Dahl, 1997, pp. 73–74).

Crookes' "torrent of charged molecules" theory is a good example of how an erroneous theoretical model can give rise to a correct experimental prediction such as the rotating radiometer. And as reported by Whittaker (1951, p. 353), Eduard Riecke showed in 1881 that when one investigates the equations of motion for such charged molecules of a given mass, a good account of the deviation of the cathode rays by a magnetic field could be deduced.

Various clouds of doubt hung over Crookes' theory, only one of which will be mentioned here.[23] Goldstein, the aether theorist, considered the issue of the mean free path of the molecules at the low vacuums Crookes was using. He determined by calculation that, for the best low vacuums that could be obtained, the mean free path of the gas molecules was about 0.6 cm according to Crookes' own model of his electrified molecules. In contrast experiment showed that the cathode rays travelled at least 90 cm, that is more than 150 times the calculated mean free path. So Crookes' "molecular torrent" model was in trouble (Dahl, 1997, p. 77; also Weinberg, 1990, p. 24).

It was Thomson who much later developed an argument against Crookes' theory that cathode rays are moving charged molecules, but only after the electron or corpuscle theory had been proposed. He argued in his 1903 book, *The Conduction of Electricity Through Gases*[24] that the momentum of the much smaller impacting charged particles would have been insufficient to cause the observed rotation of the fly. According to Thomson it is the heating of the vanes of the radiometer fly caused by the impacting particles; this generates a temperature difference which in turn produces the rotation. Such a mechanical effect supposedly due to the transfer of momentum of Crookes' molecular torrent could not have been caused by the very small corpuscles that Thomson envisaged in 1897.

Crookes played a number of variations on his view that cathode rays were a charged molecular torrent. He sometimes called the rays "charged molecules", and sometimes even "charged matter" or "radiant matter". He also referred to them in an 1879 lecture as a "fourth state of matter":

> In studying this Fourth State of Matter we seem at length to have within our grasp and obedient to our control the little indivisible particles which with good warrant are supposed to constitute the physical basis of the Universe.… We have actually touched the border land where Matter and Force seem to merge into one another, the shadowy realm, between the Known and the Unknown which for me has peculiar temptations. (Dahl, 1997, pp. 71–72)

23 Whittaker (1951, p. 353) discusses not just the problem mentioned above but also Tait's objection that an expected Doppler effect was not detectable, and Hertz's objection that he had failed to uncover any deflection of cathode rays by an electric field, something he took to support a wave theory of cathode rays. See also Dahl (1997, pp. 76–77) on these problems. The authors cited canvass possible replies to the first objection. And Thomson in his paper of October 1897 showed why Hertz's failure was to be expected, and then produced an electric field deflection at lower pressures, using it as one method of measuring the mass to charge ratio. Another of Hertz's objections to the particle theory was the emission of cathode rays through a thin film of metal at one end of a tube into the outer atmosphere; however through the work of Lenard, as will be seen in Sect. 6.3, Thomson was able to turn this to the advantage of a particle theory.

24 The relevant section was republished as Thomson and Thomson (1933, pp. 7–8).

With Crookes we now have a proliferation of names for the kind of "something" that all the scientists mentioned took themselves to be investigating. Note however that the kind names carry with them theoretical connotations. If we take these connotations seriously as part of the denotation fixer, for example we take Crookes' idea that they must be a torrent of charged molecules that have picked up their charge from the cathode and are repelled by it, then there are no such things for the names to denote. The different scientists of different theoretical persuasions could not be talking about the same things. This is one of the problematic consequences of the notion of incommensurability raised by Kuhn and others and to which the pessimistic meta-induction gives some credence.

Much more could be said of the conflict between the wave and particle models of cathode rays. But the main point of this section is that no matter how radically different the models of cathode rays might be in the ontololgies they presuppose, the conditions which enable the identification of cathode rays are independent of these theoretical models. In fact the very rival models presuppose that there are such independent identity conditions for them to be rivals of the same "something" – the cathode rays – whatever they be.

6. THOMSON AND THE IDENTIFICATION OF CATHODE RAYS OUTSIDE THE CATHODE RAY TUBE

In the story told so far, most of the properties of cathode rays are extrinsic, experimentally determined properties; little is said of their intrinsic properties.[25] But this began to change by the 1890s. Moreover it was commonly thought that cathode rays were not entities parochially confined to the behaviour of cathode ray tubes; many, wave and particle theorist alike, came to believe that they were universal constituents of matter. Thomson indicates an experimental breakthrough on this matter when he said: "So far I have only considered the behaviour of the cathode rays inside the bulb, but Lenard has been able to get these rays outside the tube" (Thomson, May 1897, p. 108). The identity conditions for cathode rays are closely tied to the features of cathode ray tubes. If they are to be identified outside such tubes then how is the identification to be made? The problem of re-identification is not commonly discussed in histories of our encounter with electrons, but it is an urgent philosophical issue for the theory of denotation and identification give here.

6.1 Thomson's experimentally based theses concerning cathode rays

Though it had already been recognised by many that cathode rays were charged (charge being an intrinsic property), a further significant intrinsic property was established by Perrin in 1895 when he showed what many had supposed, viz., the charge is negative.

[25] The intrinsic/extrinsic distinction adopted here is that discussed in Lewis (1999, Chap. 6). Lewis accepts the basic style of definition proposed by others that a thing has an intrinsic property does not entail that a further thing exists; however there are difficulties that Lewis raises that need to be overcome for a more adequate definition The suggested revisions, though important, are not relevant to this paper.

Thomson investigated this further to remove an objection raised by aether theorists, viz., that something negatively charged did arise from the cathode but this something need not be the same as cathode rays. He reports the result of his redesigned experiment in two major 1897 papers (an 30th April address referred to as Thomson, May 1897 and October 1897). The conclusion he draws cautiously in the earlier paper is that "the stream of negatively-electrified particles is an *invariable accompaniment* of the cathode rays" (May 1897, p. 7, italics added). But he says something stronger in his October 1897 paper: "the negative electrification follows the same path as the rays, and that this negative electrification is *indissolubly connected* with the cathode rays" (October 1897, p. 295, italics added). His main hypothesis in both papers is that cathode rays are (the same as) negatively electrified particles (op. cit., p. 294). The two weaker claims of invariable accompaniment or indissoluble connection follow logically from the stronger identity claim; but do not establish the identity.

In what follows, some of the main experimental findings about cathode rays in Thomson's two papers will be listed as theses (A) to (G):

Thesis (A): All cathode rays are negatively charged.

Thomson's papers are within the context of a particle model of cathode rays rather than a wave model: "The electrified particle theory has for purposes of research a great advantage over the ætherial theory, since it is definite and its consequences can be predicted; with the ætherial theory it is impossible to predict what will happen under any given circumstances, as on this theory we are dealing with hitherto unobserved phenomena in the æther, of whose laws we are ignorant" (op. cit., pp. 293–294). In discussing Perrin's experiment, and his own modification of it, Thomson says: "This experiment proves that something charged with negative electricity is shot off from the cathode, travelling at right angles to it, and that this something is deflected by a magnet" (op. cit., p. 294). Thomson's talk of a "something" in this context fits well the analysis given here in terms of definite descriptions in which the "something" is captured by the variable "x".

Further experimentally determined properties of cathode rays follow in theses (B) to (G).

Thesis (B): The m/e ratios for cathode rays converge on values between 0.3×10^{-7} and 1.5×10^{-7}.

Thesis (B) concerns a convergence, using different background theories, of values of the mass-to-charge ratio, m/e, of cathode rays. Already by 1890 Schuster had developed a theory of the motion of a charged particle in a field and had set upper and lower limits to the m/e ratio. A number of different m/e ratios for different charged particles were investigated in the 1890s both in England and Germany. Thomson's two 1897 papers use different background theories for deducing a value of the m/e ratio for cathode rays (all in the range of 0.3×10^{-7}–1.5×10^{-7}; Thomson op. cit., the tables on p. 306 and p. 309).

The first method that Thomson used to determine the m/e ratio supposed a background theory about the charged particles striking a sold body causing its temperature to rise. To establish this quantitative result he assumed background theoretical hypotheses about heat and about the motion of particles in a uniform field (op. cit., pp. 302–307; Thomson, May 1897, p. 109). The second method was quite different; he measured the amount of deflection experienced by cathode rays when they travelled the same length, first in a uniform magnetic field then in a uniform electric field, and compared their ratios (op. cit., pp. 307–310). For this he assumed a simple theory about motion of a charged massive object in a field. In both cases he helped himself to some background theory to establish the experimentally determined m/e ratios. So the determination of the m/e ratios is not theory-free. This in no way detracts from the "phenomenological" character (in the sense of Sect. 3.3) of the experimental determinations. In this context, the theory used is not explanatory; it is used to deduce the quantitative ratios in conjunction with experimental information. In addition the theories Thomson used are idealisations. At best cathode rays are not perfect satisfiers of the theories he applied to them; rather they are imperfect satisfiers.

Theses (C), (D) and (E) introduce three important independence claims. The first is:

Thesis (C): The amount of magnetic deflection of cathode rays (in a constant field) is independent of the gas in the tubes.

What Thomson, building on the work of some others, showed for a number of different gases confined in tubes or jars was that the amount of deflection is always the same, assuming that the magnetic field is the same. In this respect cathode rays behaved quite differently from canal rays. This provides the basis for an inductive inference to the conclusion: for all gases the cathode ray deflection is the same.

Thesis (D): The m/e ratio is independent of the kind of gas used in the cathode tube.

This result is a consequence of Thomson's work on the m/e ratio of cathode rays investigated in different gases in the tubes such as air, hydrogen, carbonic acid, etc. (op. cit., pp. 306–307). This result can be inductively generalised: for all gases the m/e ratio of cathode rays is independent of the gas.

Thesis (E): The m/e ratios are independent of the kind of electrode used.

This thesis (see op. cit., final section) builds on the work of Plücker, Goldstein and others all of whom used electrodes of different metals (aluminium, iron, platinum, tin, lead, copper, etc.), though they noted that the appearance of the discharge varied. By inductive generalisation one can infer thesis (E).

These three independence claims provide some of the evidence for a further inductive inference to the ubiquity of the "somethings" that are cathode rays; that is, cathode rays are a distinct kind of thing that are present in all substances and are not simply due to the peculiarities of cathode tubes.

Thesis (F): (Concerning the smallness of cathode rays): cathode rays are (1) much smaller than any other known chemical element (2) by a factor of about 1/1000.

In the October 1897 paper Thomson notes that "for the carriers of the electricity in the cathode rays m/e is very small compared with its value in electrolysis. The smallness of m/e may be due to the smallness of m of the largeness of e or a combination of these two" (op. cit., p. 310). He then cites some evidence for the smallness of m (in the absence of any direct measurement of it). The first has to do with considerations due to Lenard, on which Thomson puts great reliance in both the April and October 1897 papers. Hittorf had already noticed that cathode rays could penetrate films of metal that were opaque to light.[26] Lenard experimented with a tube that had fitted, at the end where the cathode rays impacted, a thin film of different kinds of metal, such as aluminium. He then detected cathode rays travelling from the outer side of the metal film into the surrounding atmosphere (commonly air). He noted a range of their characteristic effects such as the fluorescence they can give rise to, the fact that they can be deflected by magnetic fields, and so on.

Of interest was the distance the cathode rays travelled, or their mean free path, i.e., the distance they travel before the intensity of the rays falls by a half. For cathode rays it is about half a centimetre while that of a molecule of air is 10^{-5} cm. That is, on average cathode rays travel in air thousands of times further than the constituents of air do. From this Thomson concludes: "Thus, from Lenard's experiments on the absorption of the rays outside the tube, it follows on the hypothesis that the cathode rays are charged particles moving with high velocity[27]; that the size of the carriers must be small with the dimensions of ordinary atoms or molecules." (Thomson, May 1897, p. 108). And in the October paper he adds that the m/e ratio of cathode rays is a thousandth that of the smallest known ratio, that of the hydrogen ion in electrolysis (Thomson, October 1897, p. 310). In the quotation note the distinction Thomson makes between an experimental fact (about mean free path of cathode rays outside the tube and their relative size) and the *hypothesis* that is meant to explain this, viz., that cathode rays are charged particles. It is the result of Thesis (F) that goes a considerable way to establish the hypothesis about the ubiquitous nature of cathode rays.

Thesis (G): The distance the rays travel outside the tube is only dependent on the density of the surrounding atmosphere and not the chemical nature of the outside medium (whether air, hydrogen, sulphur dioxide, etc.), nor its physical state.

This is a further independence claim. The path of the cathode rays outside the tube depends only on the density of the medium through which they travel. This also supports Thesis (F) since cathode rays must be much smaller than the atomic elements they pass through if their mean free path is greater by an order of a thousand. Being so comparatively small, they can, so to speak, get through all the gaps that there must be in atmospheres at ordinary pressures without bumping into, or being absorbed by, the atoms of the atmosphere. This adds support to the independence claim since their mean free path does not depend on the chemical nature of the atmosphere but only physical matters such as its pressure and density.

[26] See Whitaker (1951, p. 354, n. 1) for people who investigated this phenomenon, including Lenard. This raised one apparent difficulty for the particle theory of cathode rays since it was hard to think how particles could pass through solid metal, even as thin as aluminium or gold film, if they did pass at all. Thomson held the view that nothing passed through, but the negative charge on the side of the film inside the tube (due to the presence of the cathode rays in the tube) caused a negative charge on the outer side thereby causing further cathode rays to travel outside the tube in a surrounding atmosphere (Thomson, May 1897, p. 108).

[27] Thomson had already made measurements of the velocity of cathode rays which were much lower than those for rays of light, thus casting much doubt on the aether theory of cathode rays.

There are many other background considerations that also could be introduced at this point. For example, in 1896 Zeeman and Lorentz had produced a value for m/e based on the "Zeeman effect" that was in accord with Thomson's values published in the following year, but based on different considerations from those of Thomson. The historian Isobel Falconer (1987, p. 270) suggests that, when these results were published in English in late 1896 they may have given extra impetus to Thomson's investigations into cathode rays, since Thomson's experimental interests before 1896 lay largely elsewhere. Importantly she also argues that the particle hypothesis, to be considered next, was not something that just struck Thomson at the time; he was already well acquainted with such views but may have seen in the Zeeman-Lorentz work support for this hypothesis.

6.2 A new description for fixing the denotation of "cathode rays"

All of (A) to (G) are experimental discoveries about a "something" which has already been identified by the description (¶x)Px, and on the basis of which names have been introduced. If we conjoin all of the discoveries (A) to (G) and replace any "theoretical term" such as "cathode rays" by a variable x to obtain the open sentence indicated by [(A)& … &(G)]x, then we can form a new generalised definite description (¶x)[(A)& … &(G)]x. What this says is that there is some unique kind of thing that is picked out by the description, viz., the something that satisfies [(A)& … &(G)]x (where the satisfaction is either perfect or the best, sufficiently good, imperfect satisfier). Moreover what is picked out by this new description is the same as what is picked out by (¶x)Px. And these are just cathode rays. So we can now claim:

Cathode rays = (¶x)Px = (¶x)[(A) & & (G)]x.

The first description identifies cathode rays in the parochial setting of cathode rays tubes. Moreover it is couched in terms which refer only to extrinsic relations in the cathode rays. The second description contains quite new elements which arise from discoveries about cathode rays in their parochial setting, and then outside it. It is also couched in terms which refer to some of the intrinsic features of cathode rays such as charge and mass. Moreover it contains identifying features for cathode rays which at the time obtained wide currency. One of these is the distinctive m/e ratio possessed by cathode rays but not by any other "particles" known in the physics of the time. To ensure this there was an urgent need to obtain even better values for the m/e ratio than those of Thomson. Another is the distinctive angle of deflection of cathode rays in magnetic fields. Other particles would have different angles of deflection; this would serve to differentiate one particle from another if there were several in the same field (such as in a Wilson Cloud Chamber).

6.3 Thomson's startling hypothesis

Theses (A) to (G) are experimentally determined claims about cathode rays that have been included in a new denotation fixer. But Thomson makes many other claims that have not been included because they pertain to his speculative theoretical model of cathode rays, or are employed to explain some of the experimental discoveries (A) to (G) (and others). To these hypotheses we now turn, arguing that they should have no place in any denotation fixing description because they are not about anything in the world. To include them in any denotation fixer would be to render the description denotationless.

Following on from his discussion of Lenard's result concerning the smallness of cathode rays compared with any other known particle, Thomson gives us what he calls his "startling hypothesis":

> The assumption of a state of matter more finely subdivided than the atom of an element is a somewhat startling one; but a hypothesis that would involve somewhat similar consequences – viz., that the so-called elements are compounds of some primordial element – has been put forward from time to time by various chemists. (Thomson, May 1897, p. 108)

In this context Thomson mentions a contemporary astronomer Lockyer, but also Prout who had much earlier in the nineteenth century proposed a similar hypothesis, except that Prout was mistaken in thinking that the primordial element was hydrogen. Thomson is right to call his claim a *hypothesis*, in one sense of that word. The hypothesis is intended to imply, and thus explain, claims (A) to (G); but that does not preclude the hypothesis being false. The following are seven different claims H_1 to H_7 that can be found as constituents of Thomson's "startling hypothesis"; some are consistent with modern physical theory while others are not.

Hypothesis H_1: (1) There is a primordial element of matter, much smaller in mass than that of any known atomic element, and (2) it is a constituent of all matter.

This is unexceptional, but hardly uniquely identifying. However associated with it is a further claim that is clearly false and which Thomson came to reject only well after his 1897 papers:

Hypothesis H_2: The primordial element is the *only* element out of which all matter is constituted.

Thomson then develops his explanatory hypothesis: "Let us trace the consequence of supposing that the atoms of the elements are aggregations of very small particles, all similar to one another; we shall call them corpuscles, so that the atoms of the ordinary elements are made up of corpuscles and holes, the holes being predominant." (loc. cit.) Two points can be highlighted, the first being a further hypothesis:

Hypothesis H_3: There is a predominance of holes in matter.

Thomson cites no direct experimental evidence for this, though he does use it to explain why such corpuscles have a greater mean free path than any of the atoms they comprise.

The second point concerns Thomson's introduction of a kind name "corpuscle". But it is unclear what description is to be used to fix its putative denotation. If it is claims (A)–(G) then it simply denotes cathode rays. But if the term is introduced in the context of Thomson's speculative hypothesis about the nature of cathode rays then, as will be argued shortly, it has no denotation. The hypotheses at the core of the corpuscle theory are not satisfied by anything in the world. There is an ambiguity about the term "corpuscle" that can be resolved in different ways with different consequences as to whether or not it has a denotation.

Further aspects of Thomson's broad "startling hypothesis" emerge when he continues:

> Let us suppose that at the cathode some of the molecules of the gas get split up into these corpuscles, and that these, charged with negative electricity, and moving with high velocity form the cathode rays. (Thomson, May 1897, pp. 108–109)

Two further hypotheses can be identified here. The first concerns how the cathode rays arise at the cathode and the second his core identity claim:

Hypothesis H_4: The molecules of the residual gas get torn apart at the cathode releasing some of the corpuscles to form cathode rays.

Hypothesis H_5: Cathode rays are nothing but streams of corpuscles.

Thomson's October 1897 paper expands on the startling hypothesis of the earlier May paper making clear two further speculative hypotheses concerning how the primordial corpuscles come together to form atoms. When talking of Prout's earlier anticipation of a kindred hypothesis that all matter is constituted out of hydrogen atoms, he rejects this saying that it is untenable but we can "substitute for hydrogen some unknown primordial substance X" adding that "these primordial atoms … we shall for brevity call corpuscles" (Thomson, October 1897, p. 311). But what Thomson goes on to say about the primordial corpuscles definitely shows that there are no such things.

In the second 1897 paper he reiterates H_1 and the false H_2 when he says that "we have in the cathode rays matter in a new state … in which all matter – that is, matter derived from different sources such as hydrogen, oxygen, etc. – is of one and the same kind; this matter being the substance from which all the chemical elements are made up" (Thomson, October 1897, p. 312). Thomson then develops a speculative theory about how aggregations of such primordial corpuscles would hang together in a stable configuration to form atoms. This is something that reaches back to his work in the early 1880s on how centres of repellent forces might arrange themselves in stable patterns.[28] It is part of a speculative theory, not based in experiment, concerning the vortex atom as a singularity in a uniform aether suggested earlier by William Thomson and Maxwell. The theory originates in work by Helmholz on perfect fluids in which indestructible vortices emerge that obey certain laws of rotational and translational motion.[29] The theory has an application in hydrodynamics, but its more speculative use was as a theory of how the primordial atoms that constitute all matter in the universe emerge as vortices in a universal plenum such as the aether. One suggestion Thomson

[28] Thomson (1883) is an essay on how vortex rings can form stable combinations and that "the properties of bodies may be explained by supposing matter to be collections of vortex lines in a perfect fluid that fills the universe" (op. cit., p. 1). Aspects of the theory of vortex rings last for quite some time in Thomson's thinking about his corpuscles; he devotes a whole chapter to how aspects of the vortex model might work in his informal Yale lectures of 1903; see also Thomson (1911), Chap. V.

[29] For aspects of the vortex theory see Silliman (1963) and Kragh (2001). The rudiments of the vortex atom theory are set out in Maxwell's 1875 *Encyclopaedia Britannica* article on the *Atom* reprinted in Garber et al. (1986, pp. 176–215), especially pp. 197–213.

makes is that there is a law of force of the kind envisaged by Boscovich in which at small distances the force is repulsive but at greater distances is attractive – but this involves considerable mathematical complexity owing to the number of interactions involved. As an alternative he suggests a model based on experiments concerning how different numbers of floating magnets arrange themselves in patterns of equilibrium (op. cit., pp. 313–314). This leads to Thomson's further two fundamental hypothesis about his corpuscles, one about how the primordial atoms arrange themselves and a possible second hypothesis about what these atoms are:

Hypothesis H_6: As the only constituents of nature, the corpuscles are primordial atoms (not the same as chemical elements) which arrange themselves in a law-governed way to constitute all matter including chemical elements.

Hypothesis H_7: The primordial atoms are really vortices in the aether and obey aether-vortex laws.

These seven hypotheses are the core of Thomson's speculative model of his corpuscles. If we conjoin these hypotheses and create an open sentence by deleting the theoretical term "corpuscle", viz. $(H_1)\& \ldots \&(H_7)]x$, and place a definite description operator in front, then we can form a definite description $(¶x)[(H_1)\& \ldots \&(H_7)]x$. The definite description can then be used to fix a denotation for Thomson's theoretical term "corpuscle". This corresponds to the quite general use of the Ramsey Sentence, or the Lewis–Ramsey denotation fixer, which employs all the elements of a theory rather than some restricted subset of claims associated with the theory.

Does anything perfectly satisfy the open sentence, or even play the role of being the best but imperfect satisfier? Since there is no such thing as the aether, then H_7 is false; and so the description denotes nothing. However it is possible to reject H_7 while adopting H_6. This would occur if one were to adopt the view that cathode rays are really material particles but still held the view of H_6 that such charged material particles constituted all elements and still have to come together in some way to form chemical elements. Such is one way of taking Thomson's talk of corpuscles by dropping the view that there is an aether. However it is still the case that H_2 and H_6 (with or without H_7), and following in their train a false H_5, ensure that nothing either perfectly or imperfectly satisfies the open sentence. So, the definite description denotes nothing and the term "corpuscle" fails to denote.

Not all uses of the term "corpuscle" have their reference fixed in this way. As was indicated the term could just as well have its reference fixed by the quite different denotation fixer, $(¶x)[(A)\& \ldots \&(G)]x$. In this case the term is ambiguous depending on whether its denotation is to be fixed by a theory which has several false constituent hypotheses which are part of an explanatory model, or it is to be fixed by well-established experimental claims. This locates an important ambiguity at the heart of Thomson's theory concerning whether it is about anything at all, and suggests how the ambiguity can be resolved. This is a matter often obscured by talk of concepts, such as the Thomson corpuscular concept. Ontologists wish to know: "Is the concept instantiated or not?" No clear answer is forthcoming from within the theory of concepts. But answers are forthcoming in terms of the theory of generalised descriptions used to fix denotations.

7. THE TERM "ELECTRON" AND ITS MULTIPLE INTRODUCTIONS IN PHYSICS

One of the main claims in the above is that, as far as the unobservable items of science are concerned, definite descriptions are fundamental in initially picking them out as denotata while names follow in their wake picking up as their denotata what the descriptions denote. If this is the case then it is unproblematic that the very same name can be used many times over to denote quite different unobservables. In the case of individuals the proper name "John Smith" is unproblematically ambiguous in denoting many different people. Similarly for names for scientific kinds, observable and unobservable. This is so of the term "electron". The Ancient Greek term ηλεκτρον was used to denote amber. Perhaps a Kripkean story can be told of how the name was introduced in the presence of some samples and then passed on through the Ancient Greek community.[30] The Greeks knew of the attractive powers of amber, and it was for this reason that the classically trained Elizabethan scientist William Gilbert first coined the cognate term "electric" to refer to the attractive power of amber rather than the substance amber.

George Stoney is credited with introducing the term "electron" into modern physics. From the 1870s Stoney proposed the idea that there existed a smallest unit, or atom, of electric charge involved in electrolytic processes, and in 1881 he gave an estimate of the magnitude of the charge. It was only in an 1891 address that he referred to this smallest unit using the name "electron" to denote what is picked out by the description "the smallest quantity of electricity (in electrolytic processes)" (see Dahl, 1997, pp. 205/403, n. 10–34). That the Greeks used the word ηλεκτρον to denote one kind of thing and Stoney, followed by others, used the same-sounding word "electron" to refer to another should occasion no problem; it is an ambiguity that can be removed by relativisation to languages. Stoney also believed that the electrons were permanently attached to atoms, and their oscillation gave rise to "electromagnetic stresses in the surrounding ether" (cited in Arabatzis, 2001, p. 181). But this is an additional extra belief about his electrons that purports to explain something and is not part of the description that Stoney used to introduce the term "electron".

However physicists subsequently co-opted Stoney's term "electron" to refer to two quite different kinds of thing.[31] The physicist Larmor also used the term "electron" to refer to – what? Here we need to return to the vortex ring theory that Thomson used (see Sect. 6.3). In the 1860s William Thomson proposed that atoms were vortices of motion, these being permanent, indestructible, ring-like structures capable of internal motion or vibration; they are in effect singularities in a primitive substance, a continuous and perfectly elastic fluid, the aether. In this theory neither mass nor matter nor

[30] See Kripke (1981), Lecture III, for an account of how names get introduced in some baptismal ceremony for proper names and for kinds.

[31] The story sketched draws on the work of Arabatzis (2001), Falconer (1987, 2001) and Kragh (2001) but within the context of a descriptivist account of the fixing of denotation. It will be evident that the descriptions used to introduce the term "electron" are often loaded with theory and that these cannot be readily replaced by other descriptions that lack the theory loading yet still refer to the same item (if they do refer at all).

Newtonian gravitation are primitives, but have to be "accounted for" within the vortex theory in some way. This was a view also explored by Maxwell[32] and many of his followers. As mentioned it was also a theory upon which J. J. Thomson worked in the early 1880s and, as Kragh says, "provided him with a framework of thinking" (Kragh, 2001, p. 198) that lasted even beyond the first decade of the twentieth century.

Joseph Larmor also worked within this framework trying to resolve some of the difficulties it faced. In the final section of an 1894 paper Larmor, at a suggestion of Fitzgerald, introduced the term "electron" and sometimes even spoke of "free electrons" The new denotation for the term "electron" can be reconstructed as follows: the unique kind k such that all instances of k are structural features of the aether which have a vacuous core around which is distributed a radial twist in the aetherial medium, have a permanent radial vibration which can not diminish, have a vibration and fixed amplitude and phase, have the same electric charge and the same mass, and are the universal constituents of all matter.[33] Larmor's overall view, expressed in 1895, is that "material systems are built up solely out of singular points in the ether which we have called electrons and that atoms are simply very stable collocations of revolving electrons" (cited in Falconer, 2001, p. 83).

On the point of what is or is not primitive in Larmor's ontology, his position is a little clearer in his 1900 book *Aether and Matter*:

> It is not superfluous to repeat here that the object of a gyrostatic model of the rotational ether is not to represent its actual structure, but to help us to realise that the scheme of mathematical relations which defines its activity is a legitimate conception. Matter may be and likely is a structure in the aether, but certainly aether is not a structure made of matter. This introduction of a supersensual aetherial medium, which is not the same as matter, may of course be described as leaving reality behind us; and so in fact may every result of thought be described which is more than a record of comparison of sensations. (Larmor, 1900, p. vi. Also cited, in part, in Harman, 1982, p. 102)

In the final sentence Larmor gives way to phenomenalist or empiricist or instrumentalist considerations in which "reality" is left behind; the first sentence has a slightly different emphasis in its talk of models, schemes of mathematical relations and a failure to represent. But for our purposes, the interest lies in the more realist middle sentence in which the order of ontological dependence is of matter on aether, and not of aether on matter. Even if this is not to be taken too strongly as a realist claim, it has methodological implications in that the direction of methodological analysis is from aether to matter and not conversely. This "methodological realism" is underlined when Larmor goes on to say:

[32] See Maxwell's 1875 *Encyclopaedia Britannica* article on the *Atom* reprinted in Garber et al. (1986), especially pp. 197–213.

[33] These characteristics are best set out in Arabatzis (2001, p. 183). The attribution of mass is not a primitive feature of electrons, understood as singularities in the aether, but as something for which an explanatory or reductive account needs to be given. This is not a matter that need concern us here. Issues of reduction, and especially realism about theories, would be less urgent if, as Achinstein (1991) Part II suggests, we take Maxwell and his followers to be offering theories as analogical models which downplay, in varying degrees, matters about what the world is really like, though clearly many followers of Maxwell and aether theorists took their theories realistically.

> It is incumbent upon us to recognise an aetherial substratum to matter, in so far as this proves conducive to simplicity and logical consistency in our scheme of physical relations, and helpful towards the discovery of hitherto unnoticed ones; but it would be a breach of scientific method to complicate its properties by any hypothesis, as distinct from logical development, beyond what is required for this purpose. (Larmor, 1900, pp. vii–viii)

The above fleshes out the definite description used to pick out the denotation of the term "electron" as introduced by Larmor. Such a Ramsey–Lewis denotation fixer contains a number of other theoretical terms. So fixing the denotation of "electron" can only take place in the context of fixing the denotation of other terms (such as "aether" and "radial twist", providing they have not been introduced into physical theory independently of the context of Larmor's theory). What is immediately evident is that the Larmor use of the term "electron" (to denote a structural feature of the aether) cannot not have the same denotation as Stoney's use of the term (to denote a unit of electrical charge, though Larmor's electrons do have the Stoney unit of electric charge); their denotation fixers are quite different and could not pick out the same kind of thing. So there are multiple introductions of the same term to refer to things even in quite different ontological categories. But does the Larmor term "electron" have a denotation? The verdict of the world is that there is no such thing which fits the reconstructed description given above; so there is no denotation for the Larmor term "electron". As it transpired, Larmor was developing his electron theory based in aetherial vortices just when the originator of the vortex atom, William Thomson, had doubts about it saying "'I am afraid it is not possible to explain all the properties of matter by Vortex-atom Theory alone".[34]

Not only did Fitzgerald make a suggestion to Larmor that he use the term "electron" but he made a similar suggestion to Thomson about what he could call his corpuscles. Thomson's May 1987 paper in *The Electrician* is a printing of an address given on 30 April 1897 of which Fitzgerald was aware. Fitzgerald comments on Thomson's address in a paper entitled "Dissociation of Atoms"; this appears in the same issue of *The Electrician*, and surprisingly is placed directly *before* Thomson's paper.[35] Of the several issues Fitzgerald raises about the address the main one for our purposes concerns Thomson's startling hypotheses about his corpuscles being the ultimate constituents of chemical elements. Fitzgerald "expresses the hope that Professor J. J. Thomson is quite right in his by no means impossible hypothesis". But despite this he raises some critical points about the corpuscle hypothesis, and then makes a suggestion about the existence of free electrons:

> [W]e are dealing with free electrons in these cathode rays. This is somewhat like Prof. J. J. Thomson's hypothesis, except that it does not assume the electron to be a constituent part of an atom, nor that we are dissociating atoms, nor consequently that we are on the track of the alchemists. There seems every reason to suppose that electrons can be transferred from atom to atom without at all destroying or indeed sensibly changing the characteristic properties of the atom: that in fact there is a considerable analogy between a charged sphere and an atom with an electron charge. If this be so, the question of course

[34] Cited in Silliman (1963, p. 472). Falconer (2001) also lists a number of similarities and differences between Lorenz's electrons, Larmor's electrons and Thomson's corpuscles that fleshes out much more of the story than can be done here.

[35] There is no immediate reply by Thomson to Fitzgerald's prefacing paper, though there is editorial comment (see Gooday, 2001 p. 111). However Smith (2001, p. 38) argues that Fitzgerald's comments influenced his subsequent experimentation and the topics covered in his later paper of October 1987.

> arises, how far can an electron jump in going from atom to atom? Why not the length of a cathode, say, or at least from molecule to molecule along it, or anyway in nearly straight lines along it? (Fitzgerald, 1897, p. 104)

The critical and correct, but at this time still speculative, point is the claim that atoms can retain their identity despite abandoning the idea of a Thomson corpuscle. That is, what Fitzgerald calls an electron can be free of the atom with which it has been associated and move around in various ways in cathodes, cathode tubes, and elsewhere. Such an assumption of free electrons, he says, should not lead us down the path of the alchemists who sought the one thing that could be transmuted into anything else, in which case the loss of "electrons" is taken to wrongly entail the transmutation of substances. That cathode rays are free electrons is a modification that can be made to Thomson's speculative hypotheses about his corpuscles.

Nothing is said in Fitzgerald's commentary about Larmor's account of electrons as singularities in the aether; that is well in the background and nowhere comes to the fore. Nor does Thomson mention in his May 1987 paper his more speculative Hypothesis 6 about how his corpuscles are to configure themselves to form atoms; this is a matter only raised in his later paper of October 1987. Importantly, in the second paper he does not explicitly say anything about H_7, the view that cathodes rays are really structural features of the aether (a view that Larmor, and Fitzgerald partly shared but not Thomson). It is open to the reader of Fitzgerald's paper to co-opt his term "free electrons", but not Larmor's aether theory of them, and then use the term "electron" rather than Thomson's term "corpuscle" to denote cathode rays (which are, on Thomson's "hypothesis" "charged particles moving with high velocities" (Thomson, May 1897, p. 108).

The more startling character of Thomson's hypothesis that Fitzgerald queries (because it might be taken to entail the dissociation of atoms when they lose their charged particle) is "that the atoms of all the elements are aggregations of very small particles, all similar to one another; we shall call such particles corpuscles" (loc. cit.). But then we could introduce any name on the basis of Thomson's denotation fixing description. He chose "corpuscles". Fitzgerald proposed that they be called "electrons" and that is the name that caught on in the physics community. From this point onwards, there is a sociological and historical story to be told that is well-recounted in Falconer (2001) that need not be repeated here. As Falconer expresses the complexity of what went on:

> Fitzgerald rejected the importance of corpuscles for atomic structure and shifted the context of Thomson's results to Larmor's electron theory. He ensured that the term "electron" was associated with Thomson's experimental work several years before there was full assent to Thomson's theory. That "electrons" were originally proposed as an alternative interpretation of the cathode ray results to "corpuscles" was forgotten. (Falconer, 2001, p. 86)

As is well known, Thomson resisted the use of the term "electron" to refer to the same item as his term "corpuscle", a resistance that went on to about 1915, well beyond his 1906 Nobel Prize lecture in which he did not use the term "electron" even though the lecture was entitled "Carriers of Negative Electricity".[36] What this paper adds to the

[36] Thomson's Nobel Prize lecture is reprinted as Thomson 1967. See Dahl (1997, p. 188) who emphasises Thomson's concern about distinguishing the real, material, negatively charged electron from the positive electron which, in his view, remained hypothetical. A fuller account of Thomson's resistance can be found in Falconer (2001) and Kragh (2001).

story generally told is a semantic background of denotation fixing via generalised definite descriptions. Given the different descriptions that can be culled from Thomson's papers of 1897, the more important issue is not what name to use for what the correct generalised description picks out. Rather the main issue concerns the correct description to employ and what aspect of theories of cathode rays are to be omitted from such descriptions, but nevertheless play an important role in the theoretical models of cathode rays, albeit models which are in many respects, false of what they purport to model.

8. CONTINUITY IN ONTOLOGY FROM CLASSICAL TO QUANTUM ELECTRONS

The identification of the electron is a story of ontological continuity with theory change, though not a story of name continuity. It is also a story of changing criteria of identification from Plücker's initial identification to that of Thomson's, with the same thing being identified throughout. The "something" so identified survived its several changing theoretical models, such as various kinds of aetherial wave disturbance or various kinds of particle (molecule, new subatomic particle, etc.). From the beginning of the twentieth century the electron was taken to be a charged particle obeying the classical laws of physics. A change occurred when Einstein introduced the Special Theory of Relativity; the electron now obeyed relativist dynamics. However with the various Bohr theories of the electron quite new discrete, non-classical properties were introduced, so that the electron obeyed new quantum laws. J. J. Thomson's son, G. P. Thomson, was even awarded a Nobel Prize for showing that the electron is wave-like, his experimentally determined wavelength closely agreeing with an equation derived by de Broglie. Pauli also made the quite non-classical proposal that the electron obeyed an exclusion principle: no two electrons can have the same energy state in an atom. The electron was also shown to have spin. Finally the electron has its place within both Heisenberg's matrix quantum mechanics and the Schrödinger's wave equation and more recent Quantum theories (such as QED).

Are there two (or more) electrons here, at least the classical-electron and then the quantum-electron? Or is there just the same electron we have encountered under different names, different conditions of identification and different theories or models? Bain and Norton (2001) answer "yes" to the last question – as does this paper. This final section shows how the considerations raised by Bain and Norton fit the story told here.

The first continuity they locate (Bain and Norton, 2001, pp. 453–455) is that of historically stable, intrinsic properties of electrons, stable in the sense that properties of electrons that are discovered at one point in the historical development of experiment, theory and theoretical models are kept on in later historical phases (some properties might not be kept on). Historically early properties (intrinsic or extrinsic) include: charge; Millikan's determination of the non-fractional character of the charge; the m/e ratio (though better values of this were obtained over time); the degree of deflection in a given magnetic field; and so on. Later properties include spin, the Pauli exclusion property, etc. These add to the core of growing knowledge of the properties of electrons that any theory of the electron ought to preserve. If the appeal to stable properties is to assist in the identification of electrons, then Plücker's initial denotation fixer, (¶ k)Pk, did its job of work for forty years until it was replaced by a set of identifications that enabled electrons to be located outside the context of cathode ray tubes. And this new set of identifications can have further members added to it without any change in the "something" picked out by the different identifying descriptions.

A second continuity they locate is "structure", a common feature that is preserved through changes in theory; this "is simply the smallest part of the latest theory that is able to explain the success of the earlier theories" (Bain and Norton, 2001, p. 456). There is no guarantee that there will always be such a structure, but when there is, and it is combined with historically stable properties, as is the case with the electron, then new identifying conditions can emerge. However as will be argued, even if electrons perfectly satisfy the historically stable properties, they do not perfectly satisfy these structures; at best they imperfectly satisfy them in the sense set out in Sect. 2.3. To flesh out the role that structure is to play, a theory is needed of imperfect satisfaction by an entity that is the best of a set of minimally satisfactory satisfiers.

According to Bain and Norton the structure that fills the bill is the Hamiltonian or Hamiltonian for the electron in its corr esponding theory. There are a number of different Hamiltonians (a function expressing the energy of a system in terms of momentum and position (potential energy)) according as it is embedded in one or another theory. Thus, following Bain and Norton (2001, p. 456) the Hamiltonian for the electron is

$$\mathrm{H} = (\mathbf{p} - \mathrm{e}\mathbf{A})^2/2\mathrm{m} + \mathrm{e}\varphi$$

(where **p** is the momentum, e is the charge, m is the mass of the electron and **A** and φ are the vector and scalar electromagnetic potentials). Embedding this into classical dynamics is enough for the theory that Thomson needed for his account of the deflection of his corpuscles in an electric field, or Millikan needed for his oil-drop experiment. Embedding the Hamiltonian in relativity theory produces a new equation with the additional factors above those provided by classical theory:

$$\mathrm{H} = [(\mathbf{p} - \mathrm{e}\mathbf{A}/\mathrm{c})^2\mathrm{c}^2 + \mathrm{m}^2\mathrm{c}^4]^{1/2} + \mathrm{e}\varphi$$

This introduces no new property but it does describe the behaviour of electrons by taking into account relativistic effects through additional factors.

Using these equations one can form a generalised, definite description (call this "the Hamiltonian description"), which says roughly: the unique kind of thing k such that k satisfies Hamiltonian equation H (there being a different definite description for each version of the Hamiltonian). Does the electron perfectly satisfy the above Hamiltonian descriptions, or does it only imperfectly satisfy them (in the sense of satisfaction of Sect. 2.3)? The electron cannot perfectly satisfy both; in fact it perfectly satisfies neither. But it is the best imperfect satisfier of both; but it satisfies less well the first Hamiltonian description embedded in classical theory, while it satisfies better the second Hamiltonian description embedded in relativistic theory (because the latter describes the behaviour of electrons better than the former). The difference lies in the added factors and the different functional relation between the expressions of the latter.

There are further versions of the Hamiltonian. A third form takes into account the novel property of the spin of the electron; in this case a newly discovered, but subsequently stable, property of electrons is accommodated within a new theory and yields a new Hamiltonian. A fourth version is the Hamiltonian due to Dirac; and a fifth builds on this by yielding additional terms to form a quantum field-theory account of the electron within quantum electro-dynamics (QED). And so on. Again, each of these Hamiltonians forms a Hamiltonian description. Does the electron perfectly satisfy all of these Hamiltonian descriptions? It does not, but the electron remains the best but imperfect satisfier of each of these Hamiltonian descriptions, with increasing degree of satisfaction.

The upshot is that for most of the twentieth century a case can be made for saying that (1) electrons perfectly satisfy the historically stable properties listed above; and (2) electrons satisfy imperfectly, but increasingly more accurately, a succession of descriptions built out of various Hamiltonians. Both (1) and (2) can provide identifying criteria for electrons, but (2) only within the context of the theory of identifying descriptions which allows for imperfect best satisfaction. These later descriptions build on earlier identifying descriptions which are qualitative and not quantitative in that they do not employ formulae such as the various Hamiltonians. But the succession of identifying descriptions, from the first one used by Plücker and his contemporaries to those which are based on the Hamiltonian, still manage to pick out the same entity, the electron, despite dramatic change in theory and rivalry in theory.

9. CONCLUSION

The story above argues for ontological continuity of the electron from its initial identification in the absence of any theory of the electron and via only its extrinsic properties, to later identifications through new criteria which begin to involve intrinsic properties and a succession of quite different theories, some rivalling one another. The story is told using a generalised version of Russell's theory of descriptions which is shown to be a special case of theory of the Ramsey Sentence as developed by David Lewis. To apply this version of the theory of descriptions it is necessary to draw a distinction between (a) features of theories and models of the item to be modelled, the electron, that are proposed to explain (b) the non-theoretical, experimentally determined properties or observable effects of the electron. Only the latter play a role in initially picking out and identifying the electron; if the former are included, then no story of ontological continuity can be told. This need not always be the case for entities postulated in physics and elsewhere. Sometimes the descriptions used do contain large amounts of a theory of the entity to be picked out. Two such examples are Schwarzschild's postulation of black holes (though he did not name them as such) as a development of the General Theory of Relativity shortly after it was published, and the 1930 postulation of the neutrino in the context of a theory designed to save the energy conservation principle. In such cases the existence or non-existence of such entities stands or falls with the theory used in their identification. In such cases the full generalised version of Russellian descriptions as developed by Lewis comes into play. But this is not the case for electrons; they were identified in a purely experimental context in ways which were not theory dependent. This is also the case for many other entities discovered in science but not mentioned here.

Acknowledgements

Theo Arabatzis generously provided comments on the pen-ultimate version of this paper that have proved invaluable in improving it; but I doubt if I have answered all his critical points. Earlier versions of this paper were read at various seminars and workshops from which I gathered useful comments from many people, too numerous to mention: Philosophy of Science Conference Dubrovnik, Dept HPS Cambridge, philosophy of science workshops at University of Uppsala and at University of Reading, Quadrennial International Fellows of the Pittsburgh Center for Philosophy of Science Conference held at Rytro Poland, AAHPSSS conference and NZ Division AAP conferences, and finally the conference sponsored by the Archiv Henri Poincaré at the University of Nancy.

BIBLIOGRAPHY

Achinstein, P. (1991) *Particles and Waves: Historical Essays in the Philosophy of Science*. Oxford: Oxford University Press.

Arabatzis, T. (2001) The Zeeman Effect and the Discovery of Electrons. In Buchwald and Warwick (eds.), pp. 171–194.

Arabatzis, T. (2006) *Representing Electrons*. Chicago, IL: The University of Chicago Press.

Bain, J. and Norton, J. (2001) What Should Philosophers of Science Learn from the History of the Electron? In Buchwald and Warwick (eds.), pp. 451–465.

Bishop, M. and Stich, S. (1998) The Fight to Reference, or How *Not* to Make Progress in the Philosophy of Science. *Philosophy of Science*, 65, 33–49.

Buchwald, J. and Warwick, A. (eds.) (2001) *Histories of the Electron*. Cambridge, MA: MIT.

Cartwright, N. (1983) *How the Laws of Physics Lie*. Oxford: Clarendon.

Chakravartty, A. (1998) Semirealism, *Studies in the History and Philosophy of Science*, 29, 391–408.

Dahl, Per F. (1997) *Flash of the Cathode Rays: A History of J. J. Thompson's Electron*. Bristol/Philadelphia, PA: Institute of Physics Publishing.

Demopoulos, W. (2003) On the Rational Reconstruction of our Knowledge. *British Journal for the Philosophy of Science*, 54, 371–403.

Demopoulos, W. and Friedman, M. (1985) Critical Notice: Bertrand Russell's "The Analysis of Matter": Its Historical Context and Contemporary Interest. *Philosophy of Science*, 52, 621–639.

Devitt, M. (1997) *Realism and Truth*, 2nd ed. Princeton, NJ: Princeton University Press.

Devitt, M. (2005) Scientific Realism. In F. Jackson and M. Smith (eds.) *The Oxford Handbook of Contemporary Philosophy,* chapter 26, Oxford: Oxford University Press.

Falconer, I. (1987) Corpuscles, Electrons and Cathode Rays: J. J. Thomson and the Discovery of the Electron. *British Journal for the History of Science*, 20, 241–276.

Falconer, I. (2001) Corpuscles to Electrons. In Buchwald and Warwick (eds.), pp. 77–100.

Field, H. (1973) Theory Change and Indeterminacy of Reference. *Journal of Philosophy*, 70, 462–481.

Fitzgerald, G. (May 1897) Dissociation of Atoms. *The Electrician*, 39, 103–104.

Garber, E., Brush, S., and Everitt, C. (eds.) (1986) *Maxwell on Molecules and Gases*. Cambridge, MA: MIT.

Gooday, G. (2001) The Questionable Matter of Electricity: The Reception of J. J. Thomson's "Corpuscle" Among Theorists and Technologists. In Buchwald and Warwick (eds.), pp. 101–134.
Hacking, I. (1983) *Representing and Intervening*. Cambridge: Cambridge University Press.
Harman, P. (1982) *Energy, Force and Matter: The Conceptual Development of Nineteenth Century Physics*, Cambridge: Cambridge University Press.
Kragh, H. (2001) The Electron, The Protyle, and the Unity of Matter. In Buchwald and Warwick (eds.), pp. 195–226.
Kripke, S. (1980) *Naming and Necessity*. Oxford: Blackwell.
Larmor, J. (1900) *Aether and Matter: A Development of the Dynamical relations of the Aether to Material Systems*. Cambridge: Cambridge University Press.
Laudan, L. (1981) A Confutation of Convergent Realism. *Philosophy of Science*, 48, 19–49.
Lewis, D. (1983) How to Define Theoretical Terms. In *Philosophical Papers Volume I*. Oxford: Oxford University Press.
Lewis, D. (1999) *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.
Lewis, P. (2001) Why the Pessimistic Induction is a Fallacy. *Syntheses*, 129, 371–380.
Magie, W. (1935) *A Source Book in Physics*. New York: McGraw-Hill.
Maxwell, G. (1970) Structural Realism and the Meaning of Theoretical Terms. In M. Radner and S. Winokur (eds.) *Analyses of Theories and Methods of Physics and Psychology: Minnesota Studies in the Philosophy of Science Volume IV*. Minneapolis, MN: University of Minnesota Press, pp. 181–192.
Millikan, R. A. (1965) The Electron and the Light-Quant from the Experimental Point of View. *Nobel Lectures: Physics: 1922–41*. Amsterdam: Elsevier, pp. 54–66.
Newman, M. (1928) Mr. Russell's "Causal Theory of Perception". *Mind*, 37, 137–148.
Papineau, D. (1996) Theory-Dependent Terms. *Philosophy of Science*, 63, 1–20.
Plücker, J. (1858) On the Action of the Magnet upon the Electric Discharge in Gases. *Philosophical Magazine*, 16, 119–135 and 408–418; translation by F. Guthrie of a paper of 1858 in German.
Putnam, H. (1978) *Meaning and the Moral Sciences*. London: Routledge.
Russell, B. (1927) *The Analysis of Matter*. London: George Allen and Unwin.
Russell, B. (1956) Logic and Knowledge. In R. C. Marsh (ed.) London: George Allen and Unwin.
Russell, B. (1959) *The Problems of Philosophy*. Oxford: Oxford University Press.
Shiers, G. (1974) Ferdinand Braun and the Cathode Ray Tube. *Scientific American*, 230(3), 92–101.
Silliman, R. (1963) William Thomson: Smoke Rings and Nineteenth Century Atomism, *Isis*, 54, 461–474.
Smith, G. (2001) J. J. Thomson and the Electron, 1897–1899. In Buchwald and Warwick (eds.), pp. 21–76.
Stark, J. (1964) Structural and Spectral Changes of Chemical Atoms. *Nobel Lectures: Physics: 1901–21*, Amsterdam: Elsevier, pp. 427–435.
Thomson, J. J. (1883) *A Treatise on the Motion of Vortex Rings*. London: Macmillan.
Thomson, J. J. (May 1897) Cathode Rays, Discourse Delivered at the Royal Institution, April 30. *The Electrician*, 39, 104–109.
Thomson, J. J. (October 1897) Cathode Rays. *Philosophical Magazine*, 44, 293–316.
Thomson, J. J. (1911) *Electricity and Matter*. London: Constable and Company.
Thomson, J. J. (1967) Carriers of Negative Electricity. *Nobel Lectures: Physics: 1922–41*, Amsterdam: Elsevier, pp. 145–153.
Thomson, J. and Thomson, G. (1933) *Conduction of Electricity Through Gases*, Vol. II. Cambridge: Cambridge University Press, 3rd edition (of 1903 first edition in one volume).
Weinberg, S. (1990) *The Discovery of Subatomic Particles*. New York: Freeman.
Whittaker, E. (1951) *A History of Theories of Aether and Electricity Volume 1*. London: Thomas Nelson.

SOME OPTIMISM FOR THE PESSIMIST

Commentary on "The Optimistic Meta-induction And Ontological Continuity: The Case Of The Electron", by Robert Nola

STEVE CLARKE

Robert Nola (this volume) presents a detailed case study of the history of the scientific investigation of the electron, and develops a strong argument for the referential continuity of the term "electron", in the face of a history of significant change in our favoured theories about electrons. Nola argues that we can reliably induce that future scientific theories will continue to refer to many of the same entities in the world, such as the electrons that we have discovered, even if these theories differ significantly from our current theories. This is his optimistic meta-induction (hereafter OMI). Nola's OMI is directed against proponents of the pessimistic meta-induction (hereafter PMI), such as Laudan (1981) and (particularly) Putnam (1978, pp. 24–25) who stress the number of theoretical entities, such as phlogiston, that have failed to survive the test of time and who dispute the referential continuity of apparently referentially stable theoretical terms such as "electron", in the face of significant change in our favoured scientific theories.

The PMI is one of the most potent weapons in the intellectual arsenal of opponents of scientific realism (Stanford, 2003). Scientific realists hold that the theories that mainstream contemporary scientists advocate are true or approximately true, and that the theoretical entities that are postulated in contemporary scientific theories – quarks, electrons, genes and so on – exist. Because the scientific realist is committed to the view that contemporary scientific theories are true or approximately true, the scientific realist is also committed to the view that past scientific theories that are distinct from contemporary scientific theories are false, and that the entities postulated in such theories do not exist. Because the scientific realist asks us to believe that contemporary chemical theory is true or approximately true and that oxygen exists, she is committed to the view that distinctly different past chemical theories, which once were dominant, such as Lavoisier's phlogiston theory, are false and that there is no such entity as phlogiston.

A scientific realist of Lavoisier's day would have been committed to the view that phlogiston theory is true, or approximately true, and that phlogiston exists. Thinking about the standpoint of a scientific realist of the past prompts the question posed to the contemporary scientific realist by Putnam: " how do you know you aren't in error

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 203–206.

now?" (Putnam, 1978, p. 25). If we consider the many times in the past in which the then-contemporary scientific realists would have advocated belief in the truth or approximate truth of theories that we now believe to be false and in the existence of entities that we now believe not to exist, then a meta-induction to the pessimistic conclusion that our current theories will probably turn out to be false and our current favoured theoretical entities will probably turn out not to exist (the PMI), begins to appear very compelling.

Nola urges us to reject the PMI, and to accept the rival OMI in its stead. Merely pointing out a single exception to a generalisation would not be a very good way of mounting a case for either the rejection of that generalisation or acceptance of a rival generalisation. But I take it that Nola is doing more than pointing to an exception to a generalisation. Rather, he is giving us a prescription to show how referential continuity can be established, in cases where there may be a superficially appealing argument for referential discontinuity. The payoff, if this strategy works, is that hopefully we can establish referential continuity for a core group of the significant theoretical entities of science. Not just electrons, but atoms, protons and so on. And this does look like a challenge that the advocate of the PMI must take seriously.

Advocates of the PMI might respond by taking issue with the account of reference that Nola's strategy relies on, or by disputing the extent to which this strategy could be successful, given the historical record. I leave these possibilities open for others to explore. What I want to do here, is consider whether or not an advocate of the PMI could live with the OMI, which is to ask whether the PMI is – at least in some form – compatible with the OMI.

The PMI can be something of a double-barrelled weapon, when deployed against the scientific realist. Advocates of the PMI urge a pessimistic induction regarding the referential stability of postulated theoretical entities (hereafter $PMI_{entities}$) and a complementary pessimistic induction regarding the truth-aptness of successive scientific theories (hereafter $PMI_{theories}$). The relationship between these two components of the PMI is not entirely clear, but it is clear enough that typical advocates of the PMI take the two components of the PMI to reinforce one another, so a counterexample to one component of the PMI might be thought to be problematic for the PMI as a whole.

But it seems possible that an advocate of the PMI could give up on one component of the PMI and retain acceptance of the other component. Nola's argument for the OMI, is directed against the $PMI_{entities}$. Even if it is judged to be successful (and I think it should be judged to be successful) it has no direct effect on the $PMI_{theories}$. In fact there is a way in which Nola can be understood to have strengthened the hand of the advocate of the $PMI_{theories}$. By successfully employing an account of reference that allows continuity of reference across different theories, in the case of theoretical entities, Nola acts to block the inference from referential success to theoretical success because, on his view, many theories that involve successful reference to entities have been shown to be false theories. So, in the process of undermining the $PMI_{entities}$, Nola insulates the $PMI_{theories}$ from attacks that employ arguments which appeal to considerations of reference.

But can we consistently advocate the $PMI_{theories}$ while rejecting the $PMI_{entities}$? I think the answer here is yes. To do this we can accept that ontological continuity is best explained by referential success, and so accept that the theoretical terms of science succeed in referring to theoretical entities that genuinely exist, while denying the truth-aptness of scientific theories and denying that that the accuracy of scientific theories is responsible for referential success. This is the stance of entity realists (Cartwright, 1999; Clarke, 2001). Entity realism can be understood as a position that occupies middle ground between scientific realism and empiricist antirealisms, such as the influential "constructive empiricist" antirealism advocated by van Fraassen (1980). Scientific realists argue that our best scientific theories are true or approximately true and that the entities that figure in these scientific theories exist. Constructive empiricists argue that we lack warrant for the conclusion that our best scientific theories are true or approximately true and we also lack warrant for the conclusion that theoretical entities exist. Entity realists attempt to find middle ground between these two extremes, arguing that antirealist criticisms of the claim that we are warranted in believing that our best scientific theories are true or approximately true are successful, but also arguing that, nevertheless, we are warranted in believing that the entities of contemporary science exist.

Because entity realists deny that referential success is due to theoretical success, entity realists owe us a story that explains how it is that we can come to have referential success without also having theoretical success. As those who are familiar with the work of Cartwright (1999) and Hacking (1983) will be aware, this alternative story is provided by appeal to our success at manipulating entities and intervening in their activities. We can, as Hacking (1983) famously noted, spray positrons and electrons, using them as tools to investigate the behaviour of other unobservable entities. And our ability to successfully manipulate positrons and electrons, for the purposes of experimental interventions, is sufficient warrant for acceptance of the reality of positrons and electrons, or so entity realists argue. The theoretical entities of past science that we now reject, such as phlogiston, are possible entities that we did not succeed in intervening in the activities of and which we did not succeed in manipulating. If we go along with this line of entity realist thinking we can account for referential success without appealing to theoretical success.

So, the proponent of the PMI can accept the OMI, by restricting the PMI to the theoretical level and adopting entity realism. Nola had indicated sympathies with the entity realist position in the recent past – he has previously argued that manipulability should be understood as a "mark of the real" (Nola, 2002). He would probably not find this to be an ideal result, but it is one that he has some sympathies with and may be willing to live with.

BIBLIOGRAPHY

Cartwright, N. (1999) *The Dappled World*. Cambridge: Cambridge University Press.

Clarke, S. (2001) Defensible Territory for Entity Realism. *British Journal for the Philosophy of Science*, 52, 701–722.

Hacking, I. (1983) *Representing and Intervening*. Cambridge: Cambridge University Press.

Laudan, L. (1981) A Confutation of Convergent Realism. *Philosophy of Science*, 48, 19–48.

Nola, R. (2002) Realism Through Manipulation and by Hypothesis. In S. Clarke and T. D. Lyons (eds.) *Recent Themes in the Philosophy of Science: Scientific Realism and Commonsense*. Dordrecht, The Netherlands: Kluwer, pp. 1–23.

Putnam, H. (1978) *Meaning and the Moral Sciences.* London: Routledge and Kegan Paul.

Stanford, P. K. (2003) Pyrrhic Victories for Scientific Realism. *Journal of Philosophy*, c11, 553–572.

Van Fraassen, B. C. (1980) *The Scientific Image*. Oxford: Oxford University Press.

PART 7

IS A REALIST INTERPRETATION OF QUANTUM PHYSICS POSSIBLE?

CAN WE CONSIDER QUANTUM MECHANICS TO BE A DESCRIPTION OF REALITY?

HERVÉ ZWIRN

Abstract The paper deals with the status of Quantum Mechanics as a description of reality when quantum formalism is supplemented with the decoherence mechanism. The reasons why it is often argued that Quantum Mechanics provides nothing more than a description of the appearance of reality are examined. Then, through a comparison with Relativistic Mechanics, it is showed that, were the very notion of reality not questionable, it would be possible to support the idea that it provides a description of this reality and not only of its appearance.

Keywords Decoherence, Realism, Quantum Physics, Description, Reality.

1. CLASSICAL PHYSICS

In this paper, we examine the different status of Classical Mechanics, Relativistic Mechanics and Quantum Mechanics as description of reality. We will start by Classical Mechanics and the reasons why we accept it.

In the intuitive framework of the layman, the objects we see around us are considered as existing by themselves and independently of any observer. So, the world (including objects, strengths, etc.) is considered as existing as such. This is what is called reality. It is usual to consider Classical Physics as a good description of this reality. In particular, the description given by Classical Physics is quite close to the appearance of reality and for that reason the interpretation of the classical formalism is not a problem. That means that in a naïve realist attitude framework (the intuitive framework of the layman), the world as it appears through the phenomena (I should perhaps say the world identified with the phenomena) seems to be correctly described by the classical formalism which is never in conflict with our intuitive perceptions contrary to the relativistic formalism or the quantum formalism. For the layman, who never does sophisticated experiments, Classical Physics is in perfect agreement with the real world and it is useless to raise any particular problem of interpretation. The theoretical terms of Classical Physics seem all to have an unproblematic real referent and the classical laws are easily understandable through intuitive processes.

This description is actually much too optimistic. Classical Physics is roughly made up of Newtonian Mechanics, Maxwell Electrodynamics and Boltzman Thermodynamics. Now, these three theories contain theoretical entities and laws that are neither easy to comprehend nor simple to put in correspondence with intuitive phenomena. According to the intuitive layman's description of the world, a body thrown with a non

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 209–217.

zero initial speed doesn't go on indefinitely, the existence of electromagnetic waves is not immediately obvious and the entropy of a gas is not a concept that seems to be of a direct utility in everyday life. Let's also recall Newton's reluctance to describe gravity as a force acting at a distance even if it was needed by the formalism of his theory.

So, the world description of Classical Physics does not, in the end, conform to its naïve appearance as well as we might think a priori.

However, we are so used to the description given by Classical Physics that today nobody feels uncomfortable with it. We admit easily that if we never see a body thrown with an initial speed to pursue its course indefinitely as the inertia principle says, it is because of frictional resistance impossible to reduce to zero in the real world contrary to what happens in the idealised world that the theory describes. Similarly, if we cannot directly perceive most electromagnetic waves, we know through appropriate devices how to make them appear and nobody today would deny that they exist. Their frequent use through radio and television has made them totally familiar things to us. The concept of a force acting at a distance is no more strange and even entropy has become a familiar word.

We have admitted at the end that even if literally speaking the description of the world given by Classical Physics is not strictly in agreement with its appearance, the differences are perfectly understandable: they are due either to an idealisation of the formalism or to our imperfect human means of perception. Then this description is commonly accepted and it is often considered as useless to wonder about the interpretation of classical formalism.

2. QUANTUM PHYSICS

The situation is totally different as far as Quantum Mechanics is concerned. Quantum formalism has raised a lot of debates and many different interpretations have been proposed. It is through a comparison between the status of Relativistic Mechanics and Quantum Mechanics that I would like to examine the reasons why many authors (including myself in previous works (Zwirn, 2000)) hesitate to consider that Quantum Mechanics provides an adequate description of reality (whatever this word could mean and we will see this is part of the problem).

The hard debates around quantum formalism at the beginning of Quantum Mechanics (specially the famous debate between Einstein and Bohr) are well known. At that time the arguments were essentially focused around the problem of indeterminism. Quantum Mechanics is a probabilistic theory. It provides in general only probability that such and such measured quantity has such and such value. It does not know how to predict the result of a future measure with certainty even when the initial state of the system is known perfectly. This default seemed unacceptable to Einstein who thought that a physical theory must give certain results and not probabilistic ones. Of course, Einstein was aware of the fact that even in classical physics it happens sometimes that one cannot predict the result of a future observation. But when that happens the reason why is always that one doesn't know the state of the system with enough accuracy, which is the case if there are too many different components (as for example the molecules of a gas) or if it is not possible to precisely measure the initial state. In Quantum Mechanics, this is totally different. The probabilistic side of predictions is due to the very essence of the systems and their dynamics. This was not acceptable for Einstein who thought that Quantum Mechanics formalism was not complete because it didn't allow to describe the states of the systems with enough details as to be able to do non probabilistic predictions. He was asking for a way to add some other variables in the formalism in order to obtain predictions that are sure. This attempt is called the hidden variables theory approach.[1]

It is well known that this debate reached its culminating point with the formulation of the famous Einstein–Podolski–Rosen paradox (Einstein and Podolski, 1935) (more usually known as the EPR paradox) and that the thought experiment stated by EPR was translated into testable inequalities, which state some correlations between measurements done on a pair of particles, by John Bell in 1964 (Bell, 1964). These inequalities are entailed by the principle of separability which says roughly that two systems which have interacted in the past, have no further direct and immediate influence on each other once they are spatially separated. The point is that the rules of calculation of Quantum Mechanics imply that these inequalities can sometimes be violated. This resulted in a gradual move from the problem of indeterminism to the problem of separability, a property that the macroscopic objects of our current world respect contrary to the microscopic objects described by Quantum Mechanics.

This strange behaviour joined to the fact that quantum formalism forbids in general to consider that the physical properties of systems have definite value before having been measured led many physicists to adopt a reluctant attitude towards considering Quantum Mechanics as a satisfying description of the world even though none of them could deny the validity of the theory as a prediction tool. As I mentioned earlier, many attempts have been made to build alternative theories whose formalism would allow to describe objects as having at anytime well defined properties and such that spatially separated objects could always be independent and such that non-probabilistic predictions become possible. This is a very natural attitude: what is looked after is a theory allowing us to get back to the comfortable picture of the world given by classical physics and to abandon these strange properties brought by Quantum Mechanics.

The final answer to the possibility to build such a theory has been given by a series of experiments that culminated in the beginning of the eighties. In 1982, Alain Aspect (Aspect et al., 1982) showed with the greatest accuracy that Bell's inequalities (that provides us with a testable criterion of separability) are actually violated in our world. That means that we have to abandon any hope to build a theory both describing in a correct way the real world and keeping the good old properties of Classical Physics. Every empirically adequate theory will have strange and shocking aspects.

3. DECOHERENCE

Once it has become clear that the strange features of Quantum Mechanics don't come from an imperfection of the formalism but actually from real aspects of observed phenomena, one has to make up one's mind to use it to build a picture of the world, which

[1] For an overview, see Belinfante (1973).

is not so easy. Beyond the difficulty to get back Classical Physics as limit of Quantum Mechanics when the Planck constant h tends to zero (which is due to the fact that the quantum formalism is not analytic in h), the difficulty to give interpretation of Quantum Mechanics comes mainly from what is known as the measurement problem. In the beginning of the eighties, this problem has been at least partly solved through the decoherence mechanism sometimes called the environment theory. So, I am going to explain briefly what is the measurement problem and its proposed solution.

In the quantum formalism, there are two different ways to compute how a system evolves. The first one, which is to be used when there is no observation on the system, is the Schrödinger's equation. The second one, called the wave packet reduction principle, is used when a measure is done. This could raise no difficulty if the cases in which it is the first or the second way to compute that has to be used were clearly separated. But, it happens that some experiments can lead to different points of view depending on the fact that the measurement apparatus and the observer are considered to be included in the system or not. If they are, the Schrödinger's equation must be used. If they are not, the reduction principle must be used. Both points of view are equally justified. Yet, the predictions made using the Schrödinger's equation are totally different from those made using the reduction principle. In this latter case, the predictions are in agreement with the observed facts (once the probabilistic side of them is acknowledged). However, in the former case, when the measurement apparatus and the observer are considered as included in the system, the prediction is that the apparatus (e.g. a needle on a dial) should be in a superposed state for which the needle has no definite position. Of course, such a state has never been observed. Moreover, even the observer should be in a superposed state of consciousness! For various reasons this trouble is not eliminated simply by saying that for macroscopic objects such as apparatus, one has to use necessarily the reduction principle. It would be too long here to give the details of the numerous debates around this difficulty or to present the many different solutions that have been proposed to solve it before the decoherence mechanism.[2] I would simply like to emphasize the cumbersome aspect of this problem for any direct attempt to use the quantum formalism as a description of the world. From this formalism, it emerges that it is in general impossible to think that the properties of a physical system own definite values excepted when they have just been measured. In particular, the value of the position or the value of the momentum of a particle is generally not unique. This is obviously an important difficulty since our usual world is not like that. The tables and the chairs always seem to have (and we are tempted to say have and not seem to have) a definite position and a definite speed. How is it possible to reconcile this trivial fact with the apparently opposite predictions made by Quantum Mechanics?

One solution is given by the decoherence mechanism, proposed in the eighties by Zurek (1982), following Zeh (1970), who noted that the environment in which all physical systems are plunged must be taken into account. It is not possible in this paper to present a detailed description of this mechanism but only to sketch a very brief one.

[2] See, for example, Wheeler and Zurek (1983).

As we have seen the measurement problem comes from the fact that if, as it seems legitimate to do it, we consider that the apparatus is included in the system then the apparatus is in a superposed state after the measure, i.e. after the interaction with the system on which something is measured. The decoherence mechanism prescribes to deal with a big system including not only the apparatus and the observer but also the environment. After the measurement, this big system is also in a superposed state. That is true for the environment and also for the initial system (on which the measurement is made) and the apparatus. And there is no definite value for the measured property. So what? The gain comes from the following remark: as a human observer, we cannot observe all degrees of freedom of the environment (e.g., all the positions and speeds of the molecules of the ambient air). That's impossible for us. Now, to predict what we are going to see about the apparatus (e.g., to predict the position of the needle) from the big system state, Quantum Mechanics says that we have to use a special operation (called taking the partial trace of the density operator of the big system state) which gives the state of the apparatus from the big system state. Now, it appears (apart from some subtleties that I will leave aside) that the state we obtain through this operation is a state for which the measured property has a definite value (that means that the apparatus needle has a definite position).

So, one could think that the problem is solved and that there is no difficulty left preventing to consider that Quantum Mechanics gives a good description of reality as it appears in our everyday life since it is possible, inside its framework, to get back unique values for all the apparent properties that macroscopic objects have.

4. IS THE MEASUREMENT PROBLEM SOLVED?

This position raises nonetheless some difficulties. The main one is directly linked to the decoherence mechanism and there is still a debate between the position consisting to accept that Quantum Mechanics with the decoherence mechanism provides a correct description of the world and the position pretending that it provides only a description of the appearance of the world.[3]

It is true that the decoherence mechanism gives at least an explanation for the way phenomena appear. It allows to understand why macroscopic objects never appear in superposed states. But if the decoherence mechanism is analyzed, it becomes clear that the ultimate reason for that is due to the limitation of the human possibilities of measurement. This is because we can't measure all the degrees of freedom of the environment (such a measure will require apparatus larger that the whole universe). If we were able to do it, we would see that in fact, the system, the apparatus and the environment stay in superposed states. The conclusion is that even if these systems are in superposed states, because of our human limits, we can't be aware of the superposition. Even if they are in superposed states, we perceive them as if they were in a state whose properties have defined values. This analysis has led many physicists to adopt

[3] On this subject, among many other references see Zwirn (1907, 2000), d'Espagnat (2002) and Soler (2006).

a position according to which Quantum Mechanics gives only a description of the classical appearance of the world. It allows to understand why the macroscopic world seems to be classical but it must not be interpreted as explaining how systems become classical after decoherence since they actually stay in quantum superposed states.

I would like to suggest now that perhaps it is possible to adopt a more optimistic view and to explore the way to give to Quantum Mechanics the same status as a description of reality than the one given to Relativity.

When we say that decoherence explains o*nly* the appearance of the world, we implicitly accept the idea that, not only does the world *appear* to be classical but also that it *is* classical. So "the decoherence explains only the appearance of the world" means that even though decoherence explains the appearance of the world, it explains only that, hence it doesn't explain the remaining part, that is the real nature of the world. But if we accept the fact that the world is of a quantum nature, then a theory explaining its appearance and saying that the world, though it appears as if it was classical, retains a quantum nature, is right. Whereas Classical Mechanics which explains also the appearance the world but says that the world is of a classical nature is wrong. In this case, it is fair to say that it is Classical Mechanics that explains only the appearance of the world.

Perhaps at this stage, it is interesting to ask: why is it necessary to explain the classical appearance of the world? This is because Quantum Mechanics describes the world with quantum states and that we associate with quantum states strange effects that we usually don't see. For example, the superposition of positions is something that we never see with the objects around us. This is the reason why we try to explain how, from a quantum description it is possible to reach a classical description more in agreement with what we are used to observe. But, in the case where a description, though quantum by essence, would have no observable effect different from what a classical description would give, it should be possible to accept this description as such. Put differently, if the description of the state of a system through the quantum formalism after decoherence is not classical but has no non classical observable effect, then such a state, even quantum, is acceptable as it is, and there is no need to wonder if we must or not assimilate it to a classical state. I refer to this debate because of the numerous discussions about the question whether it is legitimate or not to consider that a quantum state that has no observable effect different from a classical state can be assimilated to a classical state. The answer given by Bernard d'Espagnat and myself in previous works was a negative one. I think now that this negative answer was coming from our a priori refusal to consider that a quantum state could look like a classical state. But one lesson from decoherence is the possibility of a classical looking quantum state.

Hence, it is possible to present decoherence slightly differently: The reservation on totally accepting decoherence as giving us a correct description was due to the refusal to consider that Quantum Mechanics was giving more than a description of the appearance because according to Quantum Mechanics, the system was remaining in a quantum state. If we seriously believe (and we must do) that a quantum state can sometimes look exactly like a classical state then it is possible to cancel this reservation and to adopt the idea that decoherence is the solution of the measurement problem since it predicts the state of the system after the measure, that this state is a quantum one and nonetheless that it is in perfect agreement with any observation. The fact that this state is a quantum one is not a problem since we can't test directly its strange effects. So we are dealing not only with the appearance of empirical reality but with empirical reality itself. If it was possible for us to do the adequate experiments we would see that this reality is not identical to a classical reality much in the same way that it is possible to detect many different electromagnetic waves through apparatus even though the visible light is the only part of them we can directly see. The fact that these experiments are forever impossible to do is a secondary aspect of the question. So, perhaps the main lesson of decoherence is that some quantum states can look like classical states.

5. RELATIVISTIC PHYSICS

To reinforce my argument, I would now like to draw a parallel with the difference between the Newtonian world and the relativistic world. For low speeds, we can have the feeling that the world is Newtonian. The theory of Relativity predicts that this is not the case, but for usual speeds, it is impossible to see the difference. Strictly speaking, this is not totally true and it is much easier to measure the difference between Newtonian Physics and Relativistic Physics even at low speeds, than it is to measure the entanglement between a system and its environment. But it is only a question of degree and not of nature. It could then be possible to criticize Relativity and to pretend that it provides only a description of the appearance of the world. But nobody does that. Everybody knows that even if the world at low speeds seems to be Newtonian, it is actually relativistic and that precise measures could make the differences apparent. Hence it is agreed that relativity provides a genuine description of reality.

The parallel is clear: even if the world appears to be classical (which means only that its appearance is compatible with a classical description) it is actually quantum. What is new is that, thanks to the decoherence theory, a quantum description and a classical description are both compatible with all human possible observations. It is then possible to argue in favour of the idea that Quantum Mechanics with decoherence is a good description of reality and that it is Classical Mechanics that is only a description of the appearance of reality.

So, are we satisfied? Not really! There are still many difficulties. The first one is that, independently of the measurement problem, the very notion of reality in the usual sense is significantly manhandled by Quantum Mechanics. For example, it is forbidden to think that physical systems have definite properties when no measurement is done. Moreover, the extension of Quantum Mechanics that takes Special Relativity into account, the so called Quantum Field Theory, even says that the existence of a particle is not a well defined property. The number of particles in a state is not a fixed number either. On the top of that, non separability and entanglement between systems forbid us to think that objects are distinct entities. Contrary to the theory of Relativity which is clearly a mechanics of well identified macroscopic objects, Quantum Mechanics is strictly speaking the theory of one object: the universe as a whole. Hence, it is no more possible to keep a simple realist attitude with the idea that the world is made of many well localised objects with well defined properties and interacting through mechanical strengths (what Bernard d'Espagnat called a multitudinist view of the world (d'Espagnat, 1976)). That means that, even if some quantum states can look like classical states, it is not possible to think that reality is similar to the picture given by Classical Mechanics. And if the concept of reality becomes fuzzy it becomes less easy to consider that Quantum Mechanics gives a correct description of the real world.

So the hard objection to the acceptance of Quantum Mechanics as a description of reality could come not from the fact that Quantum Mechanics describes only the appearance of the world but from the fact that reality as it is conceived in Classical Mechanics as has no place in Quantum Mechanics. Thus, if the very idea of reality is to be retained, the reality which Quantum Mechanics is a description of is totally different from the one traditional realists rely on.

6. CONCLUSION

At this stage, the reasoning may seem puzzling. So, let me put it a bit differently. Assume that you have a discussion with a physicist who is a realist and who claims that decoherence is the solution of the measurement problem (for example Zurek (1991), in his first papers). Your first answer is to say "you are wrong: decoherence explains only the appearance of the world and not the real nature of the world". In this answer, you are implicitly assuming that the realist thinks that the appearance of the world and the nature of the world are both classical and also thinks (as many physicists do) that decoherence forces the system to be in a final classical state, what you deny. Now, as we have seen, the realist could actually think that the appearance of the world is classical but that the real nature of the world is quantum. It will be the case if his position stems from the fact that some quantum states can look like classical states. In this case, it seems fair to accept his claim that Quantum Mechanics and decoherence give a good description of the world. But a closer analyse shows that, relying on the fact that some quantum states can look like classical states, what he thinks is that even if the nature of reality is quantum, there is no significant difference with a classical reality (the situation is similar with the fact that even if we know that relativistic effects are to be taken into account, nothing prevents us to think that the relativistic reality is roughly like classical reality plus some others effects and so, that reality is described by relativity). But at this stage, comes the big trouble from the fact that this simple form of Realism is actually not compatible with Quantum Mechanics. Even if Quantum Mechanics cannot totally rule out Realism, it is compatible only with much more sophisticated forms such as the so called "Veiled Realism" of Bernard d'Espagnat or with my three levels form of Realism. In both cases, what Quantum Mechanics describes is not directly the Reality but some entity that is not Reality in itself: the "empirical reality" for d'Espagnat and the "phenomenological reality" in my conception, whereas the Reality is the "veiled reality" for d'Espagnat's conception and the third level (the "unknowable") in mine.

Your final conclusion is then that Quantum Mechanics cannot be considered to be a good description of Reality.

BIBLIOGRAPHY

Aspect, A., Grangier, P., and Roger, G. (1982) *Physical Review Letters*, 49, 91 and *Physical Review Letters*, 49, 1804.

Belinfante, F. (1973) *A Survey of Hidden Variable Theories*. Pergamon.

Bell, J. S. (1964) *Physics*, 1, 195.

Einstein, A., Podolski, B., and Rosen, N. (1935) *Physical Review*, 47, 777.

d'Espagnat, B. (1976) *Conceptual Foundations of Quantum Mechanics*. Reading, MA: Addison-Wesley.

d'Espagnat, B. (2002) *Traité de Physique et de Philosophie*. Fayard.

Soler, L. (ed.) (2006) *Philosophie de la Physique*. L'Harmattan.

Wheeler, J. A., Zurek, W. (1983) *Quantum Theory and Measurement*. Princeton, NJ: Princeton University Press.

Zeh, H. (1970) Foundations of Physics, 1, 67.

Zurek, W. (1981, 1982) *Physical Review*, D24, 1516 and *Physical Review*, D26, 1862.

Zurek, W. (1991) Decoherence and the Transition from Quantum to Classical. *Physics Today*, October, p. 36.

Zwirn, H. (1997) La décohérence est elle la solution du problème de la mesure?. In M. Bitbol, S. Laugier (eds.) *Physique et réalité*. Editions Frontières, pp. 165–176.

Zwirn H. (2000) *Les limites de la connaissance*. Odile Jacob.

COMMENTARY ON "CAN WE CONSIDER QUANTUM MECHANICS TO BE A DESCRIPTION OF REALITY?", BY HERVE ZWIRN

SOAZIG LE BIHAN

In this paper, Hervé Zwirn rightly reminds us that any attempt to recover a "classical" picture of the world through the interpretation of quantum mechanics is a dead end. The EPR-Bell adventure has taught us that no local theory, which assigns non-contextual determinate values to all observables of a system, can return all the (empirically well-confirmed) statistical predictions of quantum mechanics. Quantum mechanics cannot be interpreted as describing a "classical" world in that sense, i.e., a world constituted of independent systems that interact locally and which have determinate properties evolving deterministically in space and time.

Philosophers of physics have turned to another fruitful project, namely that of assessing to what extent quantum mechanics can be considered a fundamental theory. There are at least two necessary criteria for a theory to be considered as a candidate for a fundamental theory. A first criterion is that the theory gives a dynamical account of all physical processes and interactions in its domain. In the case of quantum mechanics, this includes measurement interactions. A second criterion is that the theory saves the phenomena. This is simply to say that a theory can be considered fundamental only if it rightly predicts what we observe. No claim is made that these criteria are sufficient for a theory to be considered fundamental. These two criteria are minimal desiderata for a theory to be a candidate fundamental theory, and a fortiori for a theory to be considered as a (structurally, approximately, etc., according to your own preferences in the debate over scientific realism) true description of the world.

It is well known that orthodox quantum mechanics does not satisfy either of the above criteria. This is because of the so-called measurement problem. Orthodox quantum mechanics features two different kinds of dynamical evolution encoded in the Schrödinger equation and the collapse postulate. The former governs all dynamical evolution except measurement, for which the latter is applied. Central problems are that there are neither necessary nor sufficient conditions for defining exactly what a measurement is, and hence, for the application of the two dynamics to a system, nor are there clear physical justifications for why measurements should be treated as a special kind of interaction. A natural move is to include measurement processes in the domain which is governed by the Schrödinger equation. Due to the linearity of the Schrödinger equation, the upshot is that an interaction between two systems, one of which is generally

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 219–222.

in a superposed state, results in a state of superposition for both systems. Now it is uncontroversial that such a superposition state *cannot* be interpreted as representing a physical system for which the values of all observables are determinate, including the pointer observable for the measurement device. This is unfortunate because this is simply not what we seem to observe at the macroscopic level. Experiments appear to have determinate outcomes. Just adopting the Schrödinger evolution is thus not enough to show how one can recover the phenomena. More has to be said.

One attempt at resolving the problem of appearances utilizes the theory of decoherence. In contrast to the several interpretations of quantum mechanics which have been developed over the last century, the theory of decoherence does not change the formalism, nor does it provide the wave function with a new philosophical interpretation. Decoherence theory provides a dynamical account for the suppression of interference, at a certain level of description, through spontaneous interaction with the environment, and this in agreement with the Schrödinger evolution. As Hervé Zwirn explains, it is rather uncontroversial that decoherence is relevant to why observable properties *appear* classical *to us*.

Along with many others, Hervé Zwirn has contented himself with this modest thesis in previous papers, but stopped short of claiming that quantum mechanics was a fundamental theory. In the present paper, by contrast, he wants to consider a stronger claim. He argues that quantum theory, equipped with decoherence, should not be denied the status of a fundamental theory on the argument that it only gives an account of the appearances. Instead, he considers the option of a strong realism about quantum theory, and argues that, in taking that option, it is possible to consider decoherence theory as giving a solution to the measurement problem. This is to say that, according to Zwirn, a realistic stance would allow one to claim that adding the formalism of decoherence theory to the formalism of quantum mechanics is sufficient to consider quantum theory a candidate fundamental theory. That decoherence accounts only for the appearances is not a problem anymore, says Zwirn, for one can hold that the world, including those physical situations that we successfully describe classically, is really quantum. Hervé Zwirn does not himself adopt such a strong realistic stance for other reasons, but rather argues that the option is viable in the case in point.

Different issues arise here. A main problem is that for a strong realistic view of a theory to be viable, such a theory has to be provided with an interpretation so that it can be a candidate for being a fundamental theory. In the case of quantum theory plus decoherence, it is simply the case that whether a realist view of the theory is taken or not, decoherence does not completely solve the measurement problem. This in turn implies that quantum theory just fails to fulfill the minimal requirement for being a fundamental theory, hence a fortiori for making the realistic view viable. That said, some interpretations are available that could help H. Zwirn defend his point. Let us explain in more detail.

That decoherence theory does not alone provide a complete solution to the measurement problem has been pointed out for years by several authors (see, e.g., Zeh, 1995, and for a recent overview, Bacciagaluppi, 2007, and references therein). Decoherence theory studies how, at the level of a system and measurement apparatus, spontaneous interaction with the environment results in the suppression of interference. As mentioned above, it does not provide any new dynamics additional to the Schrödinger equation. Hence, by linearity of the Schrödinger evolution, any enlarged system, which includes the apparatus and the environment, ends up in a superposition state. Thus, even if decoherence theory accounts for classical behavior at the level of the components, it does not suppress the problem at the level of the enlarged system. By exactly the same argument as the one used in stating the original measurement problem, a macroscopic physical system is represented by a superposition state by the theory, a state for which there is no clear and uncontroversial interpretation. To push the argument a little bit further, decoherence makes things even worse since not only the apparatus, but also the environment and finally the entire universe can end up being represented by a superposition state.

Hervé Zwirn holds that a realistic stance provides a way around the difficulty. To take a realist stance towards a theory, one has to know what the theory is about, not just what the theory says about the observables. This might be what Hervé Zwirn has in mind when he suggests that, "if the very idea of reality is to be retained, the reality which Quantum mechanics is a description of is totally different from the one traditional realists rely on." The traditional notion of reality mentioned here is roughly the notion of the world composed of independent objects to which definite properties are assigned, and which interact locally in space and time. As mentioned above, such an interpretation is not adequate for the quantum formalism. The point is that if one wants to take quantum mechanics plus decoherence as a fundamental theory one still has to provide an interpretation of the formalism. That is to say, one has to say what the world is like if the theory is true. In the case under consideration, one has to say how the universe can be described by superposition states while it appears to have determinate values for observables. Only when one has provided this question with a workable answer can she consider taking a realist view of quantum theory.

A workable answer is provided by Everettian, or many-worlds, interpretations (and indeed, Zeh (1995) and Zurek (1991), arguably among the main defenders of decoherence theory, seem to adopt such a point of view). Roughly speaking, and without entering into any detailed account of the different versions of these interpretations, the ontology associated with Everettian interpretations is a multiplicity of realities which include observers. Including the observer in the quantum description of the universe is the crucial characteristic that makes Everettian interpretations able to consistently account for the appearance that we observe determinate values for measurement outcomes. Indeed, an observer (or a copy of the observer) is associated with each component of a superposition state. Hence, at the level of components, observation of classical-like behavior is consistently accounted for. Further, decoherence seems to provide a solution for one of the biggest problems that Everettian interpretations face, the so-called preferred basis problem. Thus, adding a version of the many-worlds interpretation to the formalism of quantum mechanics and decoherence provides us with a good candidate for a fundamental theory.

This puts us in a position to raise the question of scientific realism for quantum theory, but it certainly does not by itself justify any realist stance. Further argumentation is still needed. To adopt a realist stance for a theory is to believe that the theory is (approximately) true of the world. Depending on how committed a realist one wants to be, one or several of the following items are believed to have correspondents in the world: the theoretical entities, the theoretical laws or the theoretical structures. In general, the realist accepts that the ontology associated with the interpretation of the theory (possibly or partly) corresponds to the actual inventory of the world. Now, there are other interpretations, besides decoherence plus (presumably) a version of an Everettian interpretation, which make quantum theory a good candidate for being a fundamental theory. Further, these alternative interpretations feature different ontologies than the many-worlds type. For example, Bohm-type theories feature an ontology of particles with definite positions in deterministic motion. Bohm-type theories also give an account the appearance of determinate outcomes of measurement interactions. Thus, without entering into any detailed discussion of scientific realism or underdetermination, it is simply the case that we have different interpretations available, which all make quantum theory a good candidate for being a fundamental theory, and yet are associated with ontologies that are different (and inconsistent with one another). To advocate one of these as being (in some sense) true of the world thus requires further argumentation.

To sum up, it is certainly true that decoherence has interesting features that make it worth considering. That said, decoherence theory does not alone provide a complete answer to the measurement problem, so that the quantum formalism plus decoherence still needs an interpretation to be considered as a fundamental theory. Scientific realism is neither necessary nor sufficient to get around this difficulty. More difficulties arise. On the one hand, it is unclear what it means to take a realist view of an uninterpreted formalism. On the other hand, that a theory fulfills necessary criteria for being a candidate fundamental theory does not justify taking a realist stance towards the theory. In the case under consideration, it simply so happens that there are other consistent interpretations available, associated with different ontologies. If one wants to advocate a form of scientific realism for quantum mechanics, furnished with decoherence theory, not only has she to further provide it with a consistent interpretation (presumably of Everettian type), but she also has to argue in favor of one interpretation over others.

BIBLIOGRAPHY

Bacciagaluppi, G. (2007) The Role of Decoherence in Quantum Mechanics. In Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy (Summer 2005 Edition)*, http://plato.stanford.edu/archives/sum2005/entries/qm-decoherence/>

Zeh, H. D. (1995) *Basic Concepts and Their Interpretation*. Revised edition of chapter 2 of Giulini et al. (1996). [Page numbers refer to the preprint available online, entitled ‘Decoherence: Basic Concepts and Their Interpretation’.]

Zurek, W. H. (1991) Decoherence and the Transition from Quantum to Classical. *Physics Today*, 44 (October), 36–44. [Abstract and updated (2003) version available online, under the title ‘Decoherence and the Transition from Quantum to Classical – Revisited’.]

PART 8

ONTOLOGICAL CONTINUITY: A POLICY FOR MODEL BUILDING OR AN ARGUMENT IN FAVOUR OF SCIENTIFIC REALISM?

REASONS FOR CHOOSING AMONG READINGS OF EQUIPOLLENT THEORIES

ROM HARRÉ

Abstract The paper is intended to show that the deficit in empirically based criteria for theory choice can be overcome by making use of a criterion of ontological plausibility. This move is not only defensible theoretically, but evident in the thinking of the scientific community since the seventeenth century. It is also intended to show how conceiving of science as based on iconic model building provides the resources needed to make judgments of ontological plausibility between competing theories even when these are referred to competing paradigms. The final step will be to suggest what must be done to sustain the concept of ontological plausibility when iconic modelling runs out.

Keywords theory choice, model, ontological plausibility, positivism, realism, Bohr.

The deficit in empirically based reasons for choosing among various readings of a certain class of competing theories purporting to explain the phenomena in some domain can be remedied by making use of a criterion of ontological plausibility to supplement the weakness of empirical adequacy, predictive and retrodictive success. The main poles between which alternative readings lie are Positivism (sometimes with a flavour of Pragmatism) and Realism. Should theories be taken to refer to states of the world, or are they heuristic, psychological aids to logical operations? The tendency to make realist readings of theories is not only defensible theoretically, but is evident in the thinking of the scientific community since the seventeenth century. The argument to be set out here in defence of this tendency depends on following the recent trend to think of much scientific thinking to be based on iconic model building. This will provide the resources needed to make judgments of ontological plausibility between competing theories. However, for deep theories, such as quantum mechanics, iconic modeling runs out, that is modeling based on sources drawn from classical mechanics with its roots in the ontology of material things. Niels Bohr's philosophy of physics, recently rescued from various misinterpretations, suggests what might be done to sustain the concept of ontological plausibility when iconic modeling runs out. However this move leaves realist readings equivocal.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 225–238.

1. WHY IS THERE A PROBLEM OF THEORY CHOICE?

Faced with alternative but equipollent theories, that is theories which deliver the same predictions in the same circumstances for some empirical domain, scientists do make choices between them. There are evidently reasons for choosing among such alternatives, since such choices do not seem to be just random. As long ago as 1602 Christopher Clavius pointed out that the predictive power of alternative theories cannot be among these reasons. For any data base there are indefinitely many sets of propositions from which that data base can be recovered by logical means, whether by retrodiction or prediction. Thus, any given items of data lend next to no support to any one of the alternatives. This point has been revived in recent times as the "underdetermination of theory by data". For a survey of the contemporary form of this problem see Newton-Smith (1990).

Several attempts have been made to defend the use of non-empirical criteria for theory choice, such as simplicity, elegance and so on. The criticisms of all such projects, advanced by such authors as Bunge (1963) are very powerful. Social influences have also been proposed to fill the gap between evidence and belief (Kuhn, 1970; Latour and Woolgar, 1986). No doubt these influences are part of the social psychology of the scientific community. However, whatever non-empirical reasons for choosing among readings of equipollent theories are in vogue among scientists, the upshot must be a tendency towards positivistic interpretations of science overall.

There is a world of difference between coming to believe a theory is a true description of worldly states of affairs and choosing to give it a realist reading, that is that it expresses hypotheses about worldly states of affairs.

Are there any reasons, given that predictive power (empirical adequacy) is useless, aesthetic criteria subjective and social forces irrelevant, for choosing among readings of equipollent theories on the grounds of their degree of verisimilitude as representations of aspects of the material world?

These considerations have been taken to favour a skeptical or anti-realist attitude to the verisimilitude of theories that reaches beyond the range of empirical data, that is data that, roughly speaking, are derived from perceptible states of apparatus, instruments and, in the early days of science, unaided observation. This attitude has traditionally been expressed in making a positivist or reductionist reading of theories.

There have been two skeptical arguments which have been relevant to skepticism vis-à-vis science. Both can be found prefigured in the writings of David Hume (1777 [1956]).

- Scope or inductive scepticism: there is no rational passage from empirical facts ascertained in the here and now to putative lawful generalizations (the "problem of induction").
- Depth or theory scepticism: there is no rational passage from propositions reporting relations among observables to truth claims about putative but unobservable referents of theoretical concepts (the "problem of realism").

The history of science shows clearly that scientific thinking is *regressive*. Explanations for observable phenomena[1] are created by referring to microstructures and

[1] Throughout this paper "observable" is to be understood as whatever is perceptible under a category. This is close to the Kantian concept of "phenomenon" as used, for example, by Niels Bohr.

other unobservable states of the world, or by reference to macrostructures, such as the galactic layout of the universe, which may be equally unobservable. Either way regresses lead to the presumption of the existence of beings beyond the bounds of the senses. However, the paradox of Clavius suggests that there is no rational procedure for choosing among alternative theories based on different ontological presumptions. Hence there is no way by which the relative verisimilitude of the depictions of nature implicit in rival theories could be decided. In short *all* theories should be read positivistically.

The discussion in this paper is not concerned with truth claims for theories that refer to entities, states, processes and structures that are supposed to exist beyond the reach of human powers of perception, but rather with the reasons for giving some theories a realist reading and refusing it to others. If a theory is given a realist reading the scientific community considers it worthwhile to try to construct instruments, apparatus and techniques by means of which the referents of the descriptive terms of the theory might be revealed in some way. The issue is complicated by the fact that for the most part theories are not descriptions of the world but of models or representations of aspects of the world.

The modern history of attempts to defend scientific realism as the view that we can tell which theories offer better descriptions of the unobservable processes, structures and substances supposedly responsible for the results of observations and experiments begins with the astronomical debates of the sixteenth century. The distinction between realist claims for theories and realist readings of theories was already implicit in this period. What was the problem for the astronomers?

By the end of the sixteenth century Keplerian, Copernican, Aristotelian and Ptolemaic systems had all been shown to be capable of being used to make very nearly the same astronomical forecasts. The philosophical consequences of this were explored by several philosopher/astronomers of the period, but the most relevant for this paper was the paradox of the truth status of equipollent theories (Clavius, 1602). If the step from theory to descriptions of states of the world is mediated by deduction, then for any set of true descriptions of states of the world, there are infinitely many theories from which these descriptions can be recovered be retro or predictions. Therefore, the support offered by empirical evidence that confirms predictions deduced from any of the empirically adequate theories together with local conditions, gives no support to any one theory from this set rather than any other. The consequences of Clavius's Paradox are disturbing for a scientific realist, that is for anyone who takes the astronomical hypotheses about the structure of the universe seriously. Some philosophers of the period, notably Osiander in the preface to Copernicus's great work, chose the positivist path and denied the relevance of any question as to the verisimilitude of the astronomical theories as depictions of the universe.

However, if we follow the realist path, we must admit that an empirically false theory may still be "true", that is the theory describes what may be a good representation of states and structures of the natural world, even if the deductive consequences of the theory do not match the empirical facts very well.

Empirically true theories may still be "false", that is a theory may be a poor representation of the relevant aspects of the natural world, even if it gives rise to well confirmed predictions.

The disturbing consequences just drawn presuppose scientific realism**.** We all *act as if* scientific realism is the correct account of the relation between science and the material world. Everyone (well almost everyone) now believes that AIDS is caused HIV, that earthquakes are caused by movements of tectonic plates, that the variety of species is the result of natural selection, and so on. In these and many other examples the explanans seems to involve reference to unobservables. The task for philosophers of science is to give a viable account of what we all know to be correct, namely that scientists do have reasons for choosing some theories to be given a realist reading and so not. How do we circumvent the disturbing consequences of combining Clavius's Paradox with scientific realism in such a way as to rationalize this aspect of scientific method?

The Paradox of Clavius showed that theory choice among equally empirically adequate theories must involve more than formal comparisons according to criteria such as relative simplicity. In practice scientists have recourse to a variety of criteria, but in the end the strongest reasons for a realist reading of a theory are ontological. Theories are judged for plausibility according to the kinds of beings that realist readings of alternative theories seem to presuppose. Different root ontologies are sometimes at issue, discontinuous material atoms or continuous fields of force, for example. Different versions of a root ontology sometimes clash at the level of explanatory theory. Is the energy of electromagnetic radiation propagated continuously or discontinuously?

How are the contents of explanatory theories that reach beyond the bounds of observability arrived at? How do scientists know which existential questions are decidable by advances in instrumentation and which are not? The answers to these questions can be found by studying the invention, construction and use of *models*. Theories, sets of propositions expressed in scientific discourses, are not descriptions of an independent reality, but descriptions of models or representations of that reality. This point has been argued persuasively by many authors, notably Cartwright (1989), Giere (1979) and Rothbart (1999) in recent years. The role of models as the content of theories still needs clarifying, even in the twenty-first century (Harré, 2004). The claim that theories *are* models can only be made sense of if "model" is taken in something like the sense it has in logic. This is not the sense of either "model" or "theory" in use in the scientific community.

2. MODELS IN SCIENCE

Models are representations of systems of objects. The representing relation can be of several different kinds. There are two main roles for models in relation to theory making.

1. *Logicist*: A model is a set of objects and relations by means of which a formal calculus can be interpreted, provided that the interpreted sentences are true of "states of affairs" in the domain of the model. For models in this role, theory as a formal or mathematical system is prior to any model through which the formalism of the theory can be interpreted.
2. *Iconic*: A model in the physical sciences is usually an analogue of a material system and serves as a representation of some currently unobserved aspect of the

world. A model may later attract a formal description. In this role a model as a representation of some real state of affairs is prior to any theory with which it could be described.

An iconic model, much the more important kind of model for the purposes of this argument, is related by analogy to some empirically accessible and known system which is its source, and represents some inaccessible system and unknown system which is its subject. The general pattern can be expressed as follows:

Source :: Model :: Subject

The symbol "::" represents a pattern of positive, negative and neutral analogies, that is similarities, differences and undetermined relations between the properties of each term of the modeling relation.

For example, Wegener envisioned tectonic plates floating on the plastic core of the earth as a representation of subterranean structures and processes which could be observed. The idea came by analogy from the behaviour of ice-floes moving against each other. The subject was the subterranean processes responsible for continental drift, earthquakes and so on. The source was the observable behaviour of ice-floes. The model, analogous to both subject and source was the system of tectonic plates (Hallam, 1973).

The content of either kind of scientific model is ontologically constrained by the principle that source, model and subject must all be subtypes of the same supertype (Aronson et al., 1994). As analogues of their sources and subjects iconic models can be represented by relations of similarity (positive analogy), difference (negative analogy) with some attributes of all three components untested (neutral analogy). Since anything has some similarities and some differences with anything else, the three sub-relations of analogy must be supplemented by local considerations as to which aspects carry the most weight in assessing the plausibility of a model as a verisimilitudinous representation of whatever the theory is *ultimately* about. According to this view, a theory consists of a cluster of propositions describing the relevant model. The model is, in various ways like and unlike what it represents. It follows that the theoretical discourse which describes it inherits those relations to its ultimate subject matter. At the beginning of a research programme these remain to be made determinate.

The supertype of the type-hierarchy of which source, model and subject are subtypes serves as a reference point for assessments of the ontological plausibility of a model as a representation of the actual system, when that system is observationally inaccessible (See also Chang, 2001).

Given that the balance between source, model and subject of a theory suggests that the content of the model is ontologically plausible, it would make sense to set up a research program to try to establish whether anything like the model exists. Ontological plausibility is tied in with the idea that it is good policy to take the model of a theory seriously, and to use it to guide a search for beings of the relevant type. Sometimes it turns out that the entities, structures, processes, etc. represented by the model are, or might be observationally accessible by suitable developments in technology.

The following schema shows how the existential question required by a realist reading of a theory and the relevant model is formulated, in case the entities proposed by the model are based on a source which consists of empirically accessible observable entities and states of affairs.

Syndrome of Anthrax	::	Bacterium :: Source = Occupying Army Model of cause (Pasteur and the microscope) ↓ "Eureka! There are bacteria."
Syndrome of Influenza	::	Virus :: Source = Bacterium Model of cause (San Felice and the electron microscope) ↓ "Eureka! There are viruses."

Each vertical arrow represents a successful existential confirmation.

Since the interconnected sources of the above sequence of models consist of beings of perceivable kinds, that is they are all subtypes of the supertype "living thing", the project of testing the ontological plausibility of the models reduces to the construction and use of sufficiently sophisticated sense-extending instruments, the microscopes referred in the above schema.

So long as we are dealing with theories based on iconic models neither the Kuhnian problem of the assessment of relative verisimilitude of paradigms nor the problem of distinguishing equipollent theories is difficult to solve. Do existential explorations reveal beings of the type in question or do they reveal beings that would find a place in a different type-hierarchy, grounded in a different supertype? Or would a realist reading of a theory call for the presupposition of something impossible? The ontological component of a Kuhnian paradigm just is a type-hierarchy within which the iconic models representing relatively inaccessible levels of beings are constrained.

Let me lay out the point in a somewhat different format. The kinetic theory of gases developed in the nineteenth century around an analogy between two formulae.

The law (L1) of observables: $PV = RT$.
The law (L2) of unobservables: $pv = 1/3\,nmc^2$.
L1 describes the observable behaviour of confined gases.
L2 describes the imagined behaviour of swarms of confined molecules.

The similarity between L1 and L2 suggests that gases and swarms of molecule should be located within the same type hierarchy, of which Newtonian corpuscles and the laws of their behaviour are the supertype, that the gases *are like* ensembles of molecules. The phenomenon of Brownian motion counts for physicists as an existential test, transforming the "are like" to "are".

Provided that the putative real world analogues of the model objects are in the domain of possible experience, for example, by the use of sense-extending instruments, neither the Paradox of Clavius nor Kuhnian incommensurability presents an insoluble threat to the rationality of scientific method. The choice among theories and the higher order choice among ontological type-hierarchies is, in principle, an empirical matter, and this depends on the rationality of a realist reading.

At the same time as we resolve these skeptical doubts for a certain level of scientific theorizing, we can give a clear meaning to the concept "closer to the truth". The difficulties which many authors who have discussed this concept have found insurmountable can be put down to taking "truth in science" to be a basic property of scientific discourses. This calls for an account of the relation between a proposition and a putative state of affairs. For sciences built predominantly on iconic models, such as biology and chemistry, the basic epistemic relation is between the members of the triad "source-model-subject". These are material beings, and truth of the model to its subject is assessed by a thing-to-thing comparison. Whatever the strength of that relation it will be inherited by the relevant discourse. However, since that discourse is definitive of the model, that is in relation to the model it is a set of necessary truths, there is no need to invoke a correspondence account of the truth of the description of a model. The concept "nearer to the truth" can be explicated but only for the whole complex of a discourse and its associated model.

3. MANIPULABILITY

So far the argument has turned on the possible extension of the domain of perception to reveal the presence or absence of putative beings whose type characteristics have been fixed by modeling with the relevant type-hierarchy. In the mid-seventeenth century Robert Boyle added an important extension to the argument to include the domain of potentially manipulable beings in the reach of iconic model building (Boyle, 1688).

He grounded his *methodology* on the principle that there are beings which are beyond the possibility of perception, even with sense-extending instruments, but which can be manipulated by operations on perceptible material stuff to perceptible outcomes. Boyle based his *argument* for a corpuscularian ontology (Boyle, 1688) on the principle that "Mechanical Operations have Mechanical Effects" whether the entities affected are observable or not, a principle for which he thought there was good inductive evidence. To link mechanical operations to a corpuscularian ontology he invoked the distinction between primary and secondary qualities. This distinction was made famous by Boyle's colleague at Oxford, John Locke. The primary qualities of material things can be reduced to their bulk, figure, texture (that is structure) and motion. These are perceived as they are in the world as it exists independently of human sensibility. They are the qualities of bodies with which the science of mechanics deals. The secondary qualities of material beings are their powers to induce sensory experiences such as colour, taste and warmth. According to seventeenth-century thinking they are produced in a sensitive organism by the action of some cluster of primary qualities. Heating something speeds up imperceptible molecules, which results in a feeling of warmth.

A change in a perceptible effect of a secondary quality brought about by a mechanical action on some other perceptible quality can only have come about by a manipulation of unobservable primary qualities, since these are the only real properties of material corpuscles. Boyle concluded that the existence of phenomena like the change in colour of a material substance when it is ground to powder, must be due to a change in the arrangement of "insensible parts". That is the only change grinding could produce.

This same form of argument can be seen in the presentation of the results of the Stern Gerlach experiment. Manipulation of magnetic fields reorients the nuclei of silver atoms so that they separate into two groups, depending on quantum number. When the magnet is activated the original single image produced by the ion beam splits in two, each component taking up a different orientation on the screen. The relevant Boylean principle for this experiment is that electromagnetic manipulations can have only electromagnetic effects. Thus the orientation of the image establishes the reality of electromagnetically characterized ions, since they are the only beings that could be subject to the experimental manipulations.

The general concept of "iconicity" must be enlarged to include a richer ontology. The ontology of the physical sciences can include beings accessible in principle either by devices that extend our powers of perception, or by Boylean manipulations.

Note well that the arguments above work only by virtue of the transfer of ontological plausibility from accessible material systems serving as sources for models as representations of inaccessible subjects, via the supertype that specifies the type hierarchy to which all three objects, source, subject and model, belong. Locke introduced the concept of causal power to explicate the meaning of attributions of secondary qualities, such as colour, to material things. A colour is nothing but the power a material surface has to induce a certain kind of sensation in human beings. However, his ontology was conservative. He grounded these powers in clusters of primary qualities. He believed, these were not themselves powers, but occurrent properties of matter, that is some specific layout of the bulk, figure, texture and motion of the insensible parts.

What happens when iconicity of models runs out? Since science proceeds by theoretical regresses that are read as descriptions of states, processes and so on, beyond the limits of what is currently accessible to observation, directly or indirectly, eventually this situation will occur. Must we then settle for deep Kuhnian incommensurability for theories drawing on incompatible ontologies? Is there no way for deciding between alternative realist readings of pairs of rationally inseparable equipollent theories at the deeper or grander levels of scientific explanatory regresses? Or between a realist reading on the one hand and a positivist reading on the other, as seems to have been the situation in the 1930s for quantum mechanics? In the next part of this paper I will try to develop the ontological move described above that can at least make a start with these problems as they resurface in the context of fundamental theories in physics.

4. BEYOND THE DOMAIN OF ACCESSIBILITY

There is no doubt that ontological regresses in physics and cosmology reach into domains of beings which are not representable by models located in type hierarchies rooted in supertypes drawn by abstraction from the domain of common experience, the domain of spatio-temporally distributed material things and events.

The problem situation amounts to this: the triumph of the natural sciences has been achieved by the imaginative step by step construction of ever deepening hypotheses about the causal mechanisms underlying observable phenomena. Each phase has involved a step beyond the current domain of what can be observed. Experimental science, particularly in the development of instrumentation and techniques of manipulation of unobservables, has followed this step-by-step imaginative enlargement of the human umwelt with revelations of new domains of observables and/or manipulables. Diseases are the result of the activities of microorganisms now observable in detail thanks to advances in microscopy. Chemical reactions in solution are the result of the redistribution of ions the transport of which can be followed by ingenious equipment, such as the ultramicroscope. However, from their very beginning the physical sciences outstripped technology. Outstripping technology it has also outstripped the reach of concepts at home in the world of everyday perception, and even its instrument-driven enlargement and Boylean manipulations.

What can we do? There seem to be three options:

1. Perhaps we should subside gracefully into positivism, à la Mach (1914) or maybe the smart new version proposed by Van Fraassen (1980). This counsel of despair amounts to the abandonment of the great project of physics, to understand the universe in all its depth. All scientific discourses would be read positivistically.
2. Would it be best to follow Einstein in declaring quantum mechanics incomplete, and struggle on in the search for hidden variables to restore determinism? The brightest and the best have failed to make progress along these lines. Here an unsustainable realist reading seems to have been in vogue.
3. The third way is to follow William Gilbert (1600), Roger Joseph Boscovich (1763) and Immanuel Kant (1788) in developing causal powers concepts to apply beyond the reach of perceptible experimental displays. Can a domain of nature, not so far beyond experience as to be wholly indescribable in experiential terms, nor yet bound into in-principle perceptibility by a reliance on concepts be drawn from the sources of iconic models?

If we choose the third option, which I believe is implicit in much "deep" physics, we encounter a rich collection of philosophical problems.

Recent arguments for claiming that we do have concepts with which to describe states and processes beyond the reach of iconic modeling to draw on in adjudicating choices between readings of empirically equipollent theories have involved a defence of an ontology of causal powers (Cartwright, 1989; Kistler and Gnassounou, 2005). Plausibility control through ontological considerations for realist readings of deep theories when the iconicity of substance grounded models has run out would then require a philosophical defence of ontologies that are different from those on which iconic modeling is based.

5. FARADAY'S FIELDS: THE SIMPLE CASE

Historians have detected the influence of Roger Joseph Boscovich on the field theories of Michael Faraday, mediated by the Boscovichean convictions of Humphrey Davy (Williams, 1975; Berkson, 1974). The conceptual structure of nineteenth-century field theory is straightforward. A field can be described in terms of spatially distributed dispositions of the form “If such and such conditions are met at such and such a place at a certain time, then such and such a phenomenon will occur”. As dispositional concepts are truth functions of descriptive clauses with perceptible referents, dispositions are observables, that is they are complex phenomena. While dispositional statements are true always, the dispositions they describe are manifested only occasionally, but in the right conditions they are always manifested, ceteris paribus. How then do we account for this? What is the permanent state of affairs which is responsible for the dispositions of the field, that is of the appropriate phenomenon, say acceleration of a test body, on the occurrence of the requisite conditions? The releasing of the heavy body does not cause it to fall. What does? Faraday’s answer, ultimately due to Boscovich and perhaps also influenced by Kant’s *Metaphysical Foundations of Natural Science,* again via Davy was, in our terminology, to propose a domain of causal powers, the occurrent groundings of dispositions.

The rationality of realist readings of the new field theories of the nineteenth century was rescued by the adoption of an Aristotelian ontology of potentials. The iconic models proposed for this domain, Faraday’s elastic strings and Maxwell’s rotating tubes are not plausible as candidates for existence. The forces internal to the strings and the dynamics of the Maxwellian model would need the invocation of the very same models that were supposed to account for the field phenomena in the first place. At best these models are heuristic and do not call for a realist reading. It makes no sense to propose a research program to isolate the “strings”.

Later, “energy” was proposed as pseudo-substantial model for grounding potentials. I believe that its role is wholly accounted for by the need to provide a neat metric for the powers of fields. However, to argue that case would be beyond the scope of this paper.

The main philosophical point is simple: we can know that a power exists even if we have no idea how or even if it has an occurrent grounding. It is proper then to give a realist reading to nineteenth-century field theories

6. NIELS BOHR’S ANALYSIS

Quantum mechanics has become a branch of energetics, that for all the talk of particles it is essentially a field theory (Bohr, 1958). But the simple scheme that Faraday worked out will not do. There is a problem with the phenomena. In certain circumstances whatever is “there” manifests itself in particle like ways and in other circumstances in wave like ways. How the causal powers of the world, as it exists independently of human interference, manifest themselves, depends on the choice of apparatus. Manifest phenomena, e.g., tracks on photographic plates or in cloud chambers, are attributes of complex beings, apparatus/world entities, scientifically indissoluble. Wave functions describe these beings. In Bohr’s view they are not descriptions of a hidden realm of quantum states. Thus, apparatus has a role like the Kantian schematized categories, giving form to something that is, from our point of view, inchoate. Nevertheless it would be wrong to draw positivistic conclusions from the intimate union of apparatus and world (Brock, 2003). The world has the power, in union with this or that specific kind of apparatus, to display particle-like phenomena in some circumstances and wave-like phenomena in others (The Complementarity Principle).

We need a new concept to make Bohr's (1958) Kantian insights clear. It is to hand in Gibson's concept of an "affordance", introduced to make sense of certain aspects of perception, when we can "see" what something is for or what it can do (Gibson, 1979). The relation of the concept of an affordance to the logic of dispositional concepts in general will be explored below.

Electrons are not revealed by the use of an apparatus, say a cloud chamber, but such an apparatus indissolubly integrated into the world, affords tracks. Affordances are observables and so must conform to the metaphysics of such beings. (The Correspondence Principle). Our explanatory model inserts electrons as the causes of tracks. But this is simplistic ontologically. There are no tracks and there are no interference patterns in the absence of apparatus or its equivalent in nature. To continue the Kantian theme, "electron" can be used only regulatively, not constitutively vis-à-vis tracks. In like manner "wave" can have only a regulative use vis-à-vis interference phenomena.

These considerations bear very heavily on the "how" of the model making methodology upon which all advanced sciences depend.

The importance of dispositions and causal powers in philosophy of physics is tied up with metaphysics in general. However, they are also essential, so many philosophers now believe, for setting up a defensible version of scientific realism, and so shaping important aspects of method, that is our ability to resolve the ambiguities described in the first two sections of this paper (Kistler and Gnassounou, 2005).

7. THE LOGICAL FORM OF DISPOSITIONAL ATTRIBUTIONS AND THE MEANING OF CAUSAL POWERS

We can think of dispositional attributions as pseudo-property ascriptions, effectively serving as rules of inference, about the circumstances in which a causal power is likely to be displayed. A causal powers account will also involve devices for expressing categoricity, since the power is supposed to be possessed/exist whether it is manifested or not. Such an account will also require devices for expressing relations, since powers not only have a source but also a target, that on which they act must have corresponding liabilities.

An affordance is a disposition, in which the conditionalizing component involves some human situation. Thus "a knife affords cutting" and "a cloud chamber affords tracks" are expressible in the usual conditional form of a dispositional ascription. Moreover, each is rooted in the hypothesis of a causal power, so that knives have the power to cut when in moving contact with certain substances, and cloud chambers have the power to manifest tracks when conjoined with a radioactive source and activated.

There may be no satisfactory formal *analysis* of "powers" concepts. If the concept of a causal power is a "hinge" in Wittgenstein's sense, it is the foundation on which all else depends. In this case there is not just one hinge, but a family of powers concepts, with overlapping similarities and differences among them. In such circumstances we have to rest content with various discursive presentations in a variety of dialects, that is with descriptions of how these concepts are used. Some powers concepts will be appropriate in psychology; others have a place in the language of physics. There may be no common form to all the cases of interest. There may be in this as in other central topics in philosophy, a field of concepts linked by family resemblances, and no common essence (Kistler and Gnassounou, 2005).

Starting with the vernacular there seems to be a broad but imperfect distinction between dispositions the display of which implicates the exercise of an active powerful particular and those which implicate a passive liability. Whenever an agent engenders an effect, it must "impact" a patient which suffers it.

Dispositions (Potentialities)

Active (Powers) Passive (Liabilities)

Capacities Capabilities Pronenesses Tendencies

Most of the concepts in this scheme are not only part of a field of family resemblances but they are clusters of family resemblance themselves.

The rules for the use of "powers" concepts include at least:

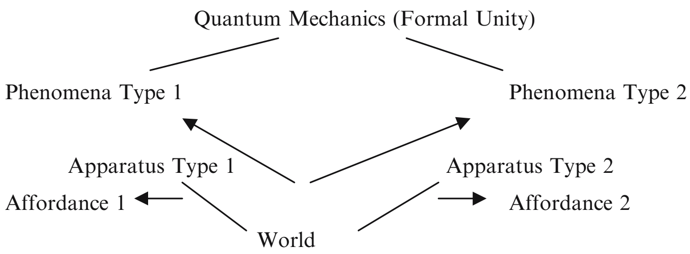
1. That there is a typical context-sensitive display of the result of the exercise of a power
2. That the relevant powerful particular pre-exists the exercise of the power on any particular occasion
3. That the display may not occur though the relevant power particular continues to exist
4. That the power can be either active or passive
5. That there may no knowable occurrent grounding of a power

"Capabilities", when activated, will typically appear in observable displays different from those of "tendencies". "Proneness" is passive power, while "Capacity" is active. (My slipping back into abstract nouns is for stylistic convenience only!) Note also that items (a) and (c) attract the possibility modality, while items (b) and (d) attract the necessity modality. Persisting causal powers are in part definitive of the natural kinds appropriate to a certain scientific discipline (Harré and Madden, 1975). If the ebonite rod no longer attracts little pieces of paper it is no longer electrically charged.

How does this link into the problems to which this paper is addressed, namely resolving the Paradox of Clavius and Kuhnian Incommensurability by adverting to the type hierarchies within which models of the unknown are locatable? The resolution of the two problems above was achieved by looking closely at how empirical evidence can be used in model-guided experimentation to resolve them both, thus vindicating realist readings of theories that refer to in-principle observables. When models represent a domain of beings inaccessible in fact but accessible in principle resolutions are not ruled out by any philosophical argument, but only by the current limitations of technology. When regresses reach beyond the range of iconic models they require the invocation of causal powers. However, because causal powers are not perceptible as such, but only in their effects, sense-extending instruments have no role here. However, physicists and engineers, chemists and pharmacists, routinely manipulate the causal powers of nature, in ways similar to Boyle's simple methodological program from the era of the dominance of the mechanistic ontology of Newtonian corpuscles.

Plainly the modeling of causal powers cannot be achieved by the source-model-subject procedure, familiar in biology, geology, chemistry and so on. It is here that mathematical representations, such as vectors and tensors, play an ineliminable part, for which there can be no iconic substitute. The development of this intuition must await another occasion.

Using the concept of an affordance we can represent Bohr's philosophy of quantum mechanics schematically.



Only combined with specific types of apparatus does the world have the power to engender phenomena. So whatever dispositions we think we can justifiably ascribe to the world must be in the form of affordances, and so ascribable only to world-apparatus complexes.

In the Bohrian scheme affordances are characterised by their dispositional manifestations. This does not open a clear and transparent window on the world, and so no plausible realist reading of physics at this level. The dispositions manifested in quantum phenomena are properties of apparatus-world complexes – not of the world *simpliciter*. The worldly causal powers cannot be uniquely characterised by any dispositional manifestations in apparatus-world complexes, unlike the Aristotelian scheme of Boscovich, Faraday and the classical field theorists. Realist readings of quantum mechanics, and its derivatives can take us no further than actual and possible apparatus-world complexes and their affordances. Such a move was made in designing the Aspect experiment in which an apparatus was devised to afford displays of indicative polarization effects.

8. CONCLUSION

A devotee of scientific realism need not despair in the face of the phenomena described in the formalism of quantum mechanics. For sure, iconic model building is inappropriate and in so far as it was tried, it led into a conceptual morass. However, the domain of perceptible *things* is not the only source for an ontology for physics. From the earliest days the alternative dynamicist metaphysics was on offer. Perhaps with the help of Bohr's insight and the conceptual developments due to Gibson a new picture of the world may emerge. More modest and more defensible realist readings will follow.

BIBLIOGRAPHY

Aronson, J. R., Harré, R., and Way, E. C. (1994) *Realism Rescued*. London: Duckworth.

Berkson, W. (1974) *Fields of Force*. London: Routledge and Kegan Paul.

Bohr, N. (1958) *Atomic Physics and Human Knowledge*. New York: Science Editions.

Boscovich, R. J. (1763) [1966] *Theoria*. Venice, published as *A Theory of Natural Philosophy*. Boston, MA: MIT.

Boyle, R. (1688) *The Origin of Forms and Qualities*. Oxford: Davis.

Brock, S. (2003) *Niels Bohr's Philosophy of Quantum Physics*. Berlin: Logos.

Bunge, M. (1963) *The Myth of Simplicity*. Englewood Cliffs, NJ: Prentice-Hall.

Cartwright, N. (1989) *Nature's Capacities and Their Measurement*. Oxford: Clarendon.

Chang, H. (2001) How to Take Realism Beyond Foot Stamping. *Philosophy*, 76, 5–30.

Clavius, C. (1602) *In Sphaeram de Johannes de Sacro Bosco*. Lyon.

Gibson, J. J. (1979) *The Ecological Approach to Visual Perception*. New York: Howard Mifflin.

Giere, R. (1979) *Understanding Scientific Reasoning*. New Tork: Holt, Rinestone & Winston.

Gilbert, W. (1600) [1958] *De Magnete*. New York: Dover.

Hallam, A. (1973) *A Revolution in Earth Sciences*. Oxford: Clarendon.

Harré, R. (2004) *Modeling: Gateway to the Unknown*. Amsterdam: Elsevier.

Harré, R. and Madden, E. H. (1975) *Causal Powers*. Oxford: Blackwell.

Hume, D. (1777) [1956] *Hume's Enquiries*, L. Selby-Bigge (ed.). Oxford: Clarendon.

Kant, I. (1788) [1970] *The Metaphysical Foundations of Natural Science*, J. Ellington (trans.). Bloomington: Bobbs Merrill.

Kistler, M. and Gnassounou, B. (2005) *Causes, Pouvoirs et Dispositions en Philosophie: Le Retour des virtus dormitivas*. Paris: Editions CNRS.

Kuhn, T. S. (1970) *The Structure of Scientific Revolutions*. Chicago, IL: Chicago University Press.

Latour, B. and Woolgar, S. (1986) *Laboratory Life*. Princeton, NJ: Princeton University Press.

Mach, E. (1914) *The Analysis of Sensations*, T. J. McCormack (trans.). Chicago, IL: Open Court.

Newton-Smith, W. (1990) *The Rationality of Science*. London: Routledge.

Rothbart, D. (1999) *Explaining Scientific Progress: Models, Metaphors and Mechanisms*. Lampeter: Mellen.

Van Fraassen, B. (1980) *The Scientific Image*. Oxford: Clarendon.

Williams, L. P. (1965) *Michael Faraday*. Englewood Cliffs, NJ: Prentice-Hall.

HARRÉ NEEDS NO REALISM

Commentary on "Reasons for Choosing Among Readings of Equipollent Theories", by Rom Harré

MAURICIO SUÁREZ[1]

It is a pleasure and an honour to comment on this paper by Rom Harré. I remember distinctly my first meeting with Rom, in May 1995, during an interview for a one-year position at Oxford. We were asked for a short presentation as part of the process, and mine was a defence of Reichenbach's views on explaining correlations by common causes. During the discussion it was objected that the logical positivists had no grounds on which to hang the distinction between accidental and law-like correlation. Rom intervened to answer the objection on my behalf: To the extent that the logical positivists had a notion of scientific law, however un-metaphysical or deflationary, they had grounds for the distinction. I was offered the position, and that was a wonderful start to my philosophical career – and a very good year for me indeed. I remember that Rom and I had lunch at Linacre several times during the year but I don't remember the issue of laws of nature coming up again – instead we talked a lot about the differences between model-theoretic and scientific models. I learnt a lot from those discussions (as from discussions with others at Oxford), but what was even more memorable and long-lasting was the optimistic feeling they aroused that my research had a definite house within Oxford "boundaries".

Rom's brief paper characteristically brings together many of the issues that have figured most prominently in his career. Three theses at least stand out, as they can be distilled from the paper, and I am in a fortunate position to comment on all of them with interest since I have defended similar theses in my recent work. First, Rom claims that ontological plausibility is an appropriate criterion to select among empirically equivalent theories; and that this criterion is indeed employed successfully by scientists. Second, Rom argues in favour of an understanding of scientific models as iconic representations. He distinguishes iconic models (i.e., the type of models one finds in science) from logicist models (i.e., the models provided by model theory). This is a distinction that we certainly discussed during my time at Oxford, and I too

[1] The writing of this paper has been supported by the Spanish Ministry of Education and Science research project number HUM2005–07187-C03–01. I thank an anonymous referee for helpful comments.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 239–243.

have defended a version in my subsequent published work. Finally, Rom has some intriguing things to say about bringing dispositional notions to bear on some of the paradoxical issues surrounding quantum theory. He claims that employing these notions can provide a satisfactory account of quantum mechanics in Bohr's Copenhagen interpretation. I too have defended a propensity interpretation of quantum mechanics in recent writings, and I agree wholeheartedly with the spirit of Rom's proposal – even though both the details and the overall philosophical framework of our respective proposals differ.

Each of these three theses is in turn employed in the paper in support of a traditional form of scientific realism, according to which mature scientific theories refer and are at least approximately true. (And indeed a defence of scientific realism has been a central focus and target of Rom's research over the years – his best known attempt is probably his joint work with Aronson and Way).[2] Thus for instance, ontological plausibility is claimed to resolve the thesis of the under-determination of theory by the data in favour of the realist position and against positivistic and instrumentalist views: "[…] In the end the strongest reasons for a realist reading of a theory are ontological. Theories are judged for plausibility according to the kinds of beings that realist readings of alternative theories seem to presuppose". A realist reading of theories also provides heuristic guidance in the application of models: "ontological plausibility is tied in with the idea that it is good policy to take the model of a theory seriously, and to use it to guide a search for beings of the relevant type". In other words, ontological plausibility is a norm of scientific research, which both guides specific inquiries and provides an abstract rule for theory-choice.

Rom's second thesis is also put to a realist use, as follows. In the iconic model of scientific representation we can distinguish between source, model and subject. (In my work I have been adopting a slightly different terminology, namely: source, model and target.) The source is some "empirically accessible and known system" which the model relates with the object of scientific research (the model's "subject") – "an inaccessible and unknown system". Some of the traditional properties of the semantic conception are then adopted to explain the relations between theories, models and real world systems, as follows.[3] The model relates to its subject non-linguistically via the holding of certain similarity relations; the theory in turn relates to the model linguistically via a description of those similarity relations in the terms of the theory within some logically structured language. Then the locution "nearer to the truth" can be explicated as a description in the language of the theory of a property of the model. (Harré defines this statement as a "necessary truth" about the model, although the notion of truth-in-the-model that seems relevant is in my view closer to that of "analytic truth".) A model is then "true" only in the sense that it is or can be (maximally) similar to its subject. The choice between models on account of their degrees of similarity to their target systems then turns into a choice between alternative theories on account of their truth content. Realism can be defended in

2 Aronson et al. (1994).

3 For the details of the semantic view see, for instance, Van Fraassen (1989, Chap. 9), and Giere (1988).

this guise, for judgements of relevant similarity between sources and subjects are putatively easier than the much more ineffable judgements regarding the truth value of theoretical statements.

Finally Rom argues that a dispositional reading of quantum mechanics enables us to interpret the theory realistically, even though the quantum domain is one where iconic modes of representation seem to falter. Of course Rom was among the very first people to think seriously about the properties described by our best science as irreducible dispositions.[4] We can apply dispositions to a Bohr-like account of quantum mechanics if we take whole complexes of quantum systems plus measurement devices to possess relational "affordances" or "capacities". This allows a scientific realist to interpret quantum mechanics coherently while insisting that it describes an independent reality – even though this is a reality at least in part made up of relational dispositional properties.

These are all intriguing theses, and I find myself in agreement with their spirit if not always their letter. But unlike Rom I am not tempted to derive the realist commitments he does from these theses. On the contrary, my response to Rom would invoke the spirit of his comment during that presentation at Oxford back in 1995. Just as the logical positivists could do fine with their notion of scientific law, however unmetaphysical or deflationary, so can we employ deflationary accounts of ontological plausibility, models, and dispositions. I have here only enough space to present a bare sketch of the relevant arguments, but in all cases the basic template of the response would be the same: to adopt the thesis, yet to then go on to insist on a deflationary or non-metaphysical reading of the concepts involved, so that the realist conclusion will not follow. In other words I would argue that Rom's three theses are not harmful to an instrumentalist provided a deflationary or non-metaphysical reading of the main entities and notions involved.

Take the first thesis. There is a reading of ontological plausibility that does not force a realist epistemology onto us; on this reading scientists choose to accept those theories that presuppose a very similar ontology to the one that we commonly uphold – or at least one that is coherent with the ontology of our ordinary life. But the fundamental reason why they will use this criterion for theory-choice has little to do with any presumed aim to capture truth in their theories. Rather it has to do with the convenience and economy afforded by the employment of a familiar system of concepts. Plausibility is always measured against a background of widely accepted beliefs – we find plausible that which exhibits reasonable connections with what we already accept.

Ontological plausibility thus refers to those existential claims that are plausible in light of other background existential assumptions. So the criterion of ontological plausibility for theory choice simply asserts that we ought to prefer theories with ontological commitments that are close to our own at present. And it seems very dubious that this can be exclusively a realist criterion for theory choice – since it would be a bizarre form of instrumentalism that requires a violation or even a reversal of ontological plausibility. For instance, an instrumentalism that requires scientists to choose

[4] Harré and Madden (1975).

those theories with as different ontological commitments as possible from our present ones seems highly implausible by instrumentalist standards themselves – for it could hardly maximise the utility of our conceptual tools to always choose theories with unfamiliar concepts and unusual existential commitments that have not been tested in the past. Similarly an instrumentalism that holds that ontological plausibility as a norm is always irrelevant to theory choice runs the risk of ignoring a basic instrumentalist insight, namely that what is relevant in each case of theory choice is determined by the purposes and peculiarities of the case itself. The instrumentalist would instead advise us to keep an open mind over whether ontological plausibility is an effective criterion for theory choice in some cases while not always being so.

What this seems to me to show is that the criterion of ontological plausibility really is neutral between realist and instrumentalist epistemologies. Both the realist and the instrumentalist can help themselves to it, and it would seem unreasonable for either to oppose the criterion as a matter of principle. Hence Rom's first thesis can be upheld without any commitment to scientific realism.

Mutatis mutandis for the second thesis: it is possible to accept that the relation between theories and the world is mediated by models and yet remain resolutely an instrumentalist about the cognitive role that theories play in enquiry.[5] I have argued that many forms of mediation by models are in fact best understood from an instrumentalist point of view, as advancing the aim of instrumental reliability rather than truth or empirical adequacy.[6] In any case nothing can prevent an instrumentalist from making all the relevant distinctions, and accepting whatever similarity orderings, without necessarily accepting the associated theories' truth – other than the truths-in-the-model that describe the properties of the model itself, including the similarity orderings. Any commitment to the ontology of the models with the higher position in the ordering seems to be an additional commitment to the source-model-subject picture, which would be naturally accepted by the realist, yet disputed by the instrumentalist.

Let us now finally turn to Rom's third thesis. I find the claim that dispositions can resolve the paradoxes of quantum mechanics both interesting and fundamentally along the right track. But I disagree with some of the details. For instance, I have argued in a different context that relational dispositional properties will not serve to solve the quantum paradoxes.[7] It is precisely here that one can find the fundamental difficulty with Popper's account.[8] Instead one needs to employ monadic propensities with relational manifestations, where the manifestations take the form of probability distributions over the values of the relevant observable. The resulting interpretation of quantum mechanics is certainly not iconic, but neither does it seem to preclude a

[5] The mediating models literature springs from (Morgan and Morrison, 1999).

[6] The key distinction is between the theory that is applied via the mediating model and the set of often diverse and even contradictory theories that are employed in extracting the commitments from the model. Since the latter theories do not always include the former one, it follows that the confirmation of the model does not necessarily constitute confirmation for the theory that the model serves to apply (See Suárez, 1999, 2005).

[7] Suárez (2004).

[8] For instance Popper (1957).

realist reading of these dispositional properties or propensities. So I agree with Rom that this is a very promising way of understanding the theory; I also agree that it *allows* for a good dose of realism. But it does not follow that realism is thereby forced upon us: realism can be allowed without being compelled. Certainly, an instrumentalist understanding of propensities and dispositional properties in general still needs to be developed. However I believe that it remains possible – and I am unconvinced by arguments that to the contrary try to exclude such an understanding on the basis of the deficiencies of the conditional statement approach to dispositions.[9]

BIBLIOGRAPHY

Aronson, J. R., Harré, R., and Way, E. (1994) *Realism Rescued*. London: Duckworth.

Bird, A. (2004) Antidotes All the Way Down. In Suárez and Bird (eds.) *Dispositions, Causes and Propensities in Science*, special issue of *Theoria*, 51, 259–269.

Giere, R. (1988) *Explaining Science*. Chicago, IL: University of Chicago Press.

Harré, R. and Madden, E. (1975) *Causal Powers*. Oxford: Blackwell.

Martin, C. (1994) Dispositions and Conditionals. *Philosophical Quarterly*, 44, 1–8.

Morgan, M. and Morrison, M. (1999) *Models as Mediators*. Cambridge: Cambridge University Press.

Popper, K. (1957) The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory. In S. Körner (ed.) *Observation and Interpretation in the Philosophy of Physics*, pp. 65–70.

Suárez, M. (1999) The Role of Models in the Application of Theories: Epistemological Implications. In Morgan and Morrison (eds.), pp. 169–198.

Suárez, M. (2004) On Quantum Propensities: Two Arguments Revisited. *Erkenntnis*, 61, 1, 1–13.

Suárez, M. (2005) The Semantic View, Empirical Adequacy and Application. *Crítica*, 37, 109, 29–63.

Van Fraassen, B. (1989) *Laws and Symmetry*. Oxford: Oxford University Press.

[9] Martin (1994) is the standard source of critiques of conditional analyses. See also Bird (2004).

PART 9

A CHANGE OF PERSPECTIVE: DISSOLVING THE INCOMMENSURABILITY PROBLEM IN THE FRAMEWORK OF A THEORETICAL PLURALISM INCORPORATING AN INSTRUMENTAL RATIONALITY

OF COURSE IDEALIZATIONS ARE INCOMMENSURABLE![1]

PAUL TELLER

Abstract Kuhnian treatment of theory comparison has engendered various worries about scientific rationality. The problems are real but, I suggest, their apparent intractability is an artifact of thinking of science as being in the exact truth business. Instead, following Cartwright, Giere, and others, think of science as being in the business of systematically developing idealized models that are always limited in scope and never completely accurate. Bear in mind that what model is appropriate depends sensitively on interests and other contextual features so that evaluation always needs to be relativized both to human interests and to the limits of human capacities. Consequently, evaluation is relative to our predictive, explanatory, and other interests, both practical and intellectual. Absent the fixed but humanly unattainable objective of truth, such subsidiary interests can appropriately vary, often as a matter of personal or collective values. Recognizing that the best that is humanly accessible are alternative idealizations, that we expect the phenomena to be usefully approached through a variety of theoretizations suddenly looks not only natural but an asset.

Keywords model, idealization, paradigm, incommensurability, theory change, truth.

1. INTRODUCTION

I propose here to address some of the problems that people have seen, in the kuhnian tradition, with theory change. My main tool will be the modeling approach to science as developed by Cartwright, Giere and others: I will suggest that the appearance of insoluble problems or radical conclusions is an artifact of developing the kuhnian insights in a framework that hangs on to thinking of science as a purveyor of exact truths rather than inexact informative models. To do this I must first prepare the ground with a sketch of the modeling view with some elaboration of my own, and likewise a sketch of prior work on trans-paradigm comparison likewise reconfigured with some of my own conclusions.

[1] Many thanks for valuable comments from Peter Lipton, Alexander Bird, Ron Giere, and others. And thanks especially to Léna Soler, Paul Hoyningen-Huene, and Howard Sankey for organizing the conference and its published proceedings that inspired and created the occasion for this work.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 247–264.

I will have already jarred readers with my use of "kuhnian" with a small "k". This is to token my policy of not attempting any analysis or proper interpretation of Kuhn himself. Instead I hope to contribute to useful ways of thinking about theory change that can be developed within the kuhnian tradition, broadly conceived, making use of insights that originated with Kuhn, but that, in countless ways, have taken on a life of their own.

2. THE MODELING APPROACH TO SCIENCE[2]

The traditional approach to science casts it in the exact truth business. This is often formulated by describing science as discovering or aiming to discover exact, universally applicable, and true natural laws that govern the operation of the "cosmic machine". Modelers instead see science as being in the model building business. Science fashions models that are always limited in scope, and even where they apply, never exactly correct.

How, on such an account, should we think of natural laws, not local, ceteris paribus or idealized generalizations, but the ones that we describe as "fundamental"? Such laws function as guidelines for building models. Laws sometimes serve this function by being incorporated into models so that they are true of the model by definition. Often they are used in the fashion of a dress pattern to be tailored to the specific case at hand for an improved fit. In either case, laws apply to the world only indirectly insofar as the model fashioned with their help is usefully similar to its target. To emphasize the idea that certain generalizations or schema function as basic model building tools, and only indirectly through the models as descriptions of the world, Giere proposes that we refer to these as (model building) *principles*.[3]

Laws, or model building principles, on this approach, are a prominent example of a great many model building tools that include also a wide range of exact and approximative mathematical techniques, model building strategies that are memorialized in a subject's exemplars, and a wide range of skills that one learns in one's training as a science apprentice. There is a great deal of affinity between the modeler's view and the kuhnian ideas of paradigms and normal science: What one learns in a science apprenticeship in normal science are the skills in applying one or another model building tool kit.

What, on such an account, is a theory? The range of models that can be built with some relatively specific model building tool kit. On this account theories have no precise boundaries. Talk of theories is an open-ended way of organizing the models that science produces along lines of what they have in common by way of principles and other tools that we use in fashioning them. This approach makes better sense than characterizing theories as the consequences of a fixed list of "postulates", better sense, for example, in describing the development and evolution of the Newtonian tradition

2 This section comprises my own summary of ideas that largely originate in the work of Cartwright, Giere, and others. See (Teller, 2001) for more summary, a historical sketch and references. For the most part I will not here give details of the origins of specific claims, analyses, and arguments.

3 See (Giere, 1999, pp. 94–95, 2006). This attitude is likewise a pervasive theme in Cartwright (1983).

from the *Principia* through the development of continuum mechanics and the methods of Hamilton and Lagrange. On this approach it is natural to expect that disciplines will mix and match the contents of various tool kits, along with the development of new tools, producing boundary cases both of specific kinds of models and whole new disciplines. This is just what we see with "semi-classical models" and disciplines such as molecular biology, physical chemistry, and chemical physics.

It is one of the fundamental tenets of the modelers' view that our models are always – in ways known and unknown – to some extent inaccurate. As such we can think of them as idealizations. I will not here discuss the many forms of idealization (and approximation) – leaving things out, keeping them in but mischaracterizing them. What I want to emphasize and what in this paper I will understand when I use "idealization" is that our models, as representations of bits of the world, are always, literally speaking, incorrect.[4]

At least in many cases, we think of our models as themselves being exact – some determinate abstract structures such as a state space or a determinate physical object serving us as a wind tunnel.[5] Insofar our models are determinate but false representations. They nonetheless function for us as representations insofar as they are similar to their targets in the respects of interest to us.

False but usefully similar. Let us work with the simplest possible case. I say that Adam is 6 ft tall. But nobody is EXACTLY 6 ft tall. So describing Adam as 6 ft tall exactly is an idealization. Nonetheless this strictly speaking false idealization works for us insofar as we use it to describe someone whose height is close enough to 6 ft tall exactly so that the discrepancy does not matter for present concerns.

We can view this example differently. Interpret the statement, "Adam is six feet tall" to mean that Adam is, not six feet exactly, but six feet CLOSE ENOUGH. We use the epithet "true" for the statement so understood just in case the discrepancy from six feet exactly does not matter for present concerns, in other worlds, in exactly those situations in which the idealization works for us as a "useful fiction". False, exact, useful fiction; or true vague statement: They get the same representational work done for us. Each is a kind of dual of the other. I will describe the two as *representational duals* or *semantic alter egos* and the claim that semantic alter egos accomplish the same representational work as *the principle of semantic alter egos*.

This is just one case. It easily generalizes to cases where the underlying discrepancy can be characterized in terms of some measurable quantity. Whether, and in what way, this idea can be further generalized, to cases such as "smart", "funny" and "mountain" is a contentious matter that I have not yet examined. Nonetheless I will here proceed on the assumption that such representational duality applies widely. The fact that the duality holds for a wide range cases in the natural sciences will suffice for present purposes.

4 As I am reminded by Ian Spencer, it is a further interesting and important question of when, and why, we dignify representations that we regard as false with the epithet "idealization". I will not pursue that question here.

5 See Teller (in preparation a) for qualification, which, however, I believe does not affect the ideas presented immediately below.

3. COMPARISON BETWEEN PARADIGMS

The next item on the bill of fare is what can be said about comparison between paradigms before we take the modeling angle into account. I will summarize some conclusions often drawn in the literature and extend them with some refinements of my own.[6] A guiding theme will be the thought that much of Kuhn's presentation was extremely idealized. Relations among paradigms are profoundly variegated and complex.

It is a common place in the literature that a great deal is carried over from one paradigm to another.[7] The move from the Newtonian tradition to both relativity and quantum mechanics saw an expansion of both instrumentation and mathematical techniques. But little was given up, and when things were abandoned, it was much more because they had become technologically obsolete rather than because they had no place in a "new paradigm". More specifically, MOST experimental results continued to have currency with the shift from, for example, the Newtonian to the relativistic tradition. It was the very same anomalous perihelion of Mercury that worried Newtonians and supported the general theory of relativity. It does not seem plausible that there was a great deal of "kuhn loss" of evidence in the change from stabilism to the plate tectonic theory, and the stabilists understood the new evidence all too clearly.[8]

Less widely appreciated is the bearing of the fact that most, if not all, of our representations are indeterminate to some extent, in some way or another. Such indeterminateness can be seen as "cushioning" what shifts of meaning do occur so that communication is nonetheless possible. To begin with, within paradigms meaning refinement are a common place. We refine by distinguishing isotopes and by distinguishing bacteria from viruses. Usage employing the less refined vocabulary is generally reconstructable from the point of view of the more refined vocabulary. There is no reason why such meaning refinement cannot occur between paradigms. Rather than seeing the meaning of "mass" as having changed so that communication is impossible we should take usage to have usefully distinguished between rest and relativistic mass, as well as between gravitational and inertial mass.[9] As in the case of intraparadigm refinement, prior usage is often easy to reconstruct from the point of view of the refined usage.

More generally, I suggest that in important cases the new is better described, not as replacing, but as incorporating and generalizing the old. It is a commonplace that Newtonian mechanics can be seen as a limiting case of special relativity as c goes

[6] I want to acknowledge my debt to Hoyningen-Huene (1993) and Bird (2000). I will not document in any detail either what I have taken from them nor the points on which I diverge, Consequently, except where I have included specific citations, they bear no responsibility for what I say here.

[7] For example, Hoyningen-Huene (1993, p. 220), McMullin (1983), Forster (2000, pp. 243–244), Bird (2000, pp. 99–122, 130–133).

[8] Constancy of evidence may hold much less well in cases such as the change from the phlogiston theory to modern chemistry, but even in this case there would appear to have been a great deal of agreement on what to count as cases of combustion and calcination, however differently such cases might have been interpreted.

[9] Earman and Friedman (1973), Field (1973), Van Fraassen (2000, p. 113). The latter refinement, of course, took place long before relativity, in which the referents are reidentified in the equivalence principle, though I would think that we would still want to count the concepts as distinct. Friedman (2000) argues that there are nonetheless profound differences due to changes in "constitutive principles" in different frameworks. See also n. 11.

to infinity. Newtonian mechanics also bears a variety of more complicated limiting relations with both general relativity and quantum mechanics. Whenever there is a natural way to see a limiting case as a special case of a more general framework this relation can be reversed so that we can see the more general framework as a kind of generalization of the limiting case.[10]

Why, then, has the kuhnian tradition often rejected the idea of generalization in favor of replacement? I suspect that there are two kinds of reasons. First, in ascending to relativity or quantum theory we discover ways in which Newtonian mechanics is FALSE. As false, we must see it as replaced, don't we? No – not as soon as we appreciate that quantum and general relativistic mechanics are themselves incomplete and not completely accurate frameworks, for a number of reasons, the most clear cut of which is probably that neither can be formulated so as to accommodate the phenomena covered by the other. But then, if general relativity and quantum theories are themselves false, in whatever sense Newtonian mechanics is false, the fact that Newtonian mechanics is false is no longer a good reason for thinking of it as rejected rather than generalized.

The second source I suspect of resistance for thinking of the new theories as generalizing the old is skepticism as to whether the mockup of Newtonian theory within the new can be taken to be exactly the old Newtonian theory. This is especially plausible for general relativity and quantum theories where the limiting relations are complex. But to be cogent, this worry must assume that there is something quite determinate that counts as THE old Newtonian mechanics. And, what, please, is this? There are two reasons why I expect to hear no good response. First, even looking at any one narrowly constrained time period, one would be hard put in giving a very precise characterization of just what constitutes Newtonian mechanics. Certainly, in the kuhnian tradition, citing the three laws of motion, or this plus the law of gravitation leaves out much too much. Second, and putting aside failure for there to be anything that would count as the completely determinate theory at any one time, there is a long and complex evolution of theory from Newton through the development of continuum and rational mechanics. I suggest that we see a continuation of this history in the further changes that may occur with reconfiguration of Newtonian concepts as limiting cases within relativistic and quantum mechanical accounts. There need be no more difficulty in comparing results between a version of Newtonian mechanics presented as, for example, a limiting case of relativity, special or general, and prior versions than in comparing results between the Newtonian mechanics of Hamilton or Lagrange and that of the *Principia*. The kinds of changes are different, to be sure, in all of these cases; but I see no reason that any one of these kinds of evolution of Newtonian ideas is more a barrier to communication and comparison than any of the others.[11]

I conclude that in the important case of Newtonian mechanics there is a useful sense in which the new can be seen as preserving rather than replacing the old, preserving the old as a limiting case. In such cases the metaphor of a "revolution" is

[10] Bohr enshrined these sorts of considerations in his "correspondence principle".

[11] Again, see Friedman (2001) for disagreement with the foregoing on many points. I expect that the prima facie differences between us will be considerably relieved by seeing his "constitutive principles" as model building requirements within a tradition, along the lines of Giere's notion of "principles". More careful consideration of these differences will have to wait for a future occasion.

misplaced. Rather than seeing such paradigms as being overthrown or abandoned we should describe them as being expanded. For these reasons, to the extent that the foregoing provides an apt description, worries about communication should disappear.

On the other hand there certainly are important cases in which we should describe the new as largely replacing the old. Kuhn uses the example of the shift from Aristotelian to Newtonian physics. My impression is that there was an even more fundamental and radical break in the whole enterprise of understanding natural phenomena in the shift from Aristotelian and Scholastic precepts to those of the new mechanical philosophy.[12] In this case my model of subsumption of paradigms does not apply at all. The advocates of the new rejected the old as sterile, empty, even unintelligible. The advocates on both sides saw themselves as having the uniquely correct way of understanding natural phenomena. The methodologies and presumed underlying metaphysics would appear to have been incommensurable in the sense of there being no "common measure" in terms of which they could be usefully compared (as opposed simply to being disparaged). In such circumstances adherents may well have been swayed by extra-rational considerations.

I choose the example of the shift to the new mechanical philosophy because, in terms of the issue of comparison between frameworks, it contrasts maximally with the case of expansion of the Newtonian framework in the development of relativistic and quantum theories. But the example is troubled as a candidate case of paradigm replacement in as much as the mechanical philosophy does not seem like a proper example of a kuhnian paradigm or disciplinary matrix. It is too schematic, not yet sufficiently spelled out as a guide to the practice of normal science. The kuhnian tradition might better classify it as preparadigmatic.

However in another way the advent of the mechanical philosophy accurately illustrates a more general point that I want to develop. The prior point about paradigm expansion rather than replacement illustrates the general consideration that it is a drastic simplification to think of the mature sciences as comprising a collection of, perhaps interrelated, but discrete paradigms or disciplinary matrices. The relation of paradigmatic elements is much more complicated than that. The mechanical philosophy illuminates this point in a different way. The mechanical philosophy, while hardly itself an articulated disciplinary matrix, did provide a powerful framework that guided the articulation of many more specific disciplinary matrices, helping us to see things that these have in common. Special relativity likewise operates as a very broad and open-ended framework that can be applied in many areas. The requirement that a more specific theory be Lorentz covariant constrains quantum field theory and functions as a requirement for the characterization of the infinitesimal in general relativity. There are, of course, enormous differences in the cases of special relativity and the mechanical philosophy; but both illustrate the point that paradigmatic elements operate at a multitude of levels of generality.

What are paradigms or disciplinary matrices? I conclude that the kuhnian tradition has gone astray insofar as it has tried to approach this subject by thinking in terms of

[12] Kuhn touches on this issue in (1957) in ways that I feel are largely consistent with the sketch I am suggesting here. See Chap. 7, "The New Universe", especially pp. 261 ff.

discrete units. The metaphor of a paradigm, both in its narrow and broad construction, has been enormously useful in helping better to understand the enterprise of science. But it is better applied as a tool in doing the natural history of a complex interleaving of elements that are organized by relations of hierarchy, overlap, borrowing, imitation, as well as, in important cases, replacement. Here the modelers' idea of a model building tool kit provides an invaluable supplementary metaphor. In most, if not all, the tool kits used in physics one will find the calculus, Galilean idealization, conservation of momentum and energy, and our present day much generalized idea of a "mechanism" that has grown out of the earlier precepts of the mechanical philosophy. The principle of Lorentz covariance and perturbational methods are widely used. Newton's laws of motion or a Schrödinger equation have more specialized, but still extremely broad application. All these examples are taken from physics, but pretty clearly similar relations will be found in chemistry and biology. Most likely there is more diversity in the social sciences, but in these subjects also the advice to look for a natural history with a very complex structure is likely to provide a sound methodology.

4. THE PROBLEM OF RADICAL INCOMMENSURABILITY

My treatment of paradigm expansion instead of replacement applies to many cases to address questions about retrospective comparison of paradigms and relative evaluation of the old in terms of the new. But this treatment has no application where the old really was abandoned. And nothing I have said so far touches the issues about prospective paradigm change. How does an adherent of a paradigm understand a new proposal and rationally evaluate it? Or is there really no conceptual room for such activities when paradigm change is substantial? Many have seen in the kuhnian tradition some kind of radical incommensurability that makes paradigm choice non-rational. There has also been a great deal of work in an effort to clarify what this notion of radical incommensurability might be and whether, as clarified, there are real problems of rational comparison. While I am in agreement with substantial parts of the critiques of the claimed irrationality, I feel that prior discussions neglect an important underlying source of the felt tension. To explain what I take this tension to be I will give my own schematic sketch of what the problem is supposed to be in the kuhnian tradition, a sketch that will make it easy to point out the source of tension that, I believe, has so far gone unnoticed.

To see how the problem arises, for the moment put aside the material outlined above on science as a model building enterprise and get yourself back into the frame of mind (if you ever left it) of thinking of science as being in the exact truth business.

It has often been noted that there is a strong kantian flavor to the kuhnian idea of a paradigm being constitutive of a world-as-known.[13] The relevant kantian attitude starts from the idea that cognitive contact with the world must be with the world

[13] Friedman (2001, pp. 30 ff.) notes that this way of thinking goes back to Reichenbach and before. See also Friedman (2001, pp. 41–46), Hoyningen-Huene (1993, pp. 31 ff.), Lipton (2001), Bird (2000, pp. 123–130), and Kuhn himself (1993, pp. 331–332). Hacking's (1993) development might be seen as having some affinity to this approach.

as conceptualized. For Kant there was also a second ingredient, that there is only one possible cognitive perspective[14] in terms of which one can structure a "world-as-known", or a system of phenomena in an extended kantian sense. As long as one includes both of these kantian doctrines there need be no special problem about truth. But kuhnians apply the first doctrine, that a world-as-known is structured through the conceptual scheme that we apply, but reject the second: They hold that, in paradigm change, we change the relevant conceptual scheme. Thus distinct paradigms give rise to distinct worlds-as-known. Furthermore, kuhnians insists that in some sense these alternative worlds-as-known compete. If we then try to understand each of these worlds-as-known in terms of the familiar idea of exact truth it would appear to follow that we have competing "exact truths" each of which, moreover, can be evaluated only within the world of the paradigm from which it came.[15]

However one might understand "radical incommensurability", the foregoing ought to qualify.

To see in another way what the problem is supposed to be let me sketch an example. In Newtonian mechanics gravitation is a force. In general relativity gravitational phenomena are an artifact of the curvature of space-time: Gravitating objects follow geodesics in a curved spatio-temporal manifold. In the one world-as-known there are gravitating objects pushed around by forces to follow non-inertial paths in a flat space-time. In the other gravitating objects, not subjected to any forces at all, move inertially along geodesics in a curved space-time. Each of these pictures is thought of as "true" as seen from within its own framework, and there is no neutral point of view from which these perspectives could be adjudicated.

In particular, kuhnians reject the idea that we should favor the perspective of general relativity because it is, somehow, closer to the truth than that of the Newtonian tradition. Apparently the argument is that evaluations for closeness to the truth can only be made within a paradigm as there is no such thing as a trans-paradigmatic perspective from which to made such comparisons; and since such comparisons in principle cannot be made, it is concluded that they make no sense.[16]

How, then, are we to come to terms with the persistent conviction that, after all, there is some sense in which our theories have "made progress"? Kuhn himself has proposed that such progress can be understood in terms of comparisons of "puzzle solving power" as measured by standards such as accuracy, consistency, breadth of scope, simplicity, and fruitfulness.[17] Insofar as such comparisons are understood in any objective, context independent sense, in particular as any comparisons with respect to truth or any context independent measure of "closeness to the truth", I join others in seeing such suggestions as unhelpful.[18] With the possible exception of consistency, all of Kuhn's

[14] When one examines Kant's full corpus this characterization may not be accurate, but it is often the way the *First Critique* is read.

[15] See Friedman (2001, pp. 47–48) who likewise attributes "[t]he underlying source of this post-Kuhnian predicament" to "the breakdown of the original Kantian conception of the a priori. (Friedman's description of the predicament does not correspond in detail to mine.)

[16] See Hoyningen-Huene (1993, pp. 263–264).

[17] Kuhn (1996, pp. 185, 206; 1970, pp. 265–266; 1977, pp. 321 ff.), Hoyningen-Huene (1993, pp. 149 ff).

[18] For example, Friedman (2001, pp. 50–53), Bird (2000, Chap. 6).

proposed trans-paradigm standards are highly subjective, so it is hard to see how they could provide the desired foundation of any objective notion of scientific progress.[19]

These problems have often polarized interpreters, some concluding that Kuhn must be entirely wrong, others concluding that we must embrace one form or another of social constructivism or irrationalism. However, there is a way to retain many of the kuhnian insights without abandoning ourselves to any problematic sort of subjectivism or relativism.

5. RADICAL INCOMMENSURABILITY AND KUHNIAN PARADIGMS AS MODEL BUILDING SYSTEMS

Two comments will point the way out of this tangle. To begin with, Kuhn himself was wary of the use of the term "true" (1996, pp. 170, 206–207) allowing its use, if at all, only intra-paradigmatically (1970, pp. 264, 266).[20] But even if we so use "true", ALL the candidates we have for fundamental natural laws are recognized not to be exactly true, and so, strictly speaking, false. Is the velocity of light constant? Yes, but only in a perfect vacuum, and there are no perfect vacua. Both quantum theories and general relativity are compromised because neither can be made to fit with the other. Conservation principles hold, strictly, only in the face of exact symmetries, which never hold exactly in the messy real world. And so on.[21]

But if our current science fails, at least at the fundamental level, to deliver exact truths, it is common to insist that our current theories, surely, are "closer to the truth" than their predecessors. As mentioned above, kuhnians reject the idea that science is moving "closer to the truth" as it moves from paradigm to paradigm.[22]

In fact I agree with kuhnians that it makes no sense to make absolute, context-independent, trans-paradigmatic comparisons of closeness to the truth, but for reasons that underwrite rather than problematize the trans-paradigmatic comparisons that we evidently do make. I think that the idea of closeness to the truth does make sense, but only in a contextualized, interest-relative way that, of course, makes it inapplicable to

[19] Kuhn himself (see the references n. 17) clearly did not see these considerations as context independent and objective, describing their application as value laden, inasmuch as (a) they are imprecise and (b) these desiderata may conflict so that the application of values are required in making choices about how to balance them. Agreed. But in addition, there is a great deal of subjectivity in these criteria individually. For example, fruitful in what respects? The more to be puzzled about how such criteria can function in clarifying the notion of objective progress in science.

[20] In (2000, p. 99) Kuhn comments that "something like a redundancy theory of truth is badly needed to replace" a correspondence theory of truth. On the next page he writes that "In this reformulation, to declare a statement a candidate for true/false is to accept it as a counter in a language game whose rules forbid asserting both a statement and its contrary." These pages suggest a strongly paradigm relative epistemic conception of "truth".

[21] Here I make this claim only for the laws of 'fundamental physics". In (in preparation, a) I argue for this kind of conclusion much more generally. Scientific claims, and claims generally, are never both completely precise and exactly true.

[22] See, for example, Kuhn (1970, pp. 265–266) where the argument I summarized in the last section occurs, together with a hint of something like the argument that I give immediately below. See also Hoyningen-Huene (1993, pp. 262–264) for discussion of this material.

the proposed job of adjudicating between paradigms if that job is understood as requiring an objective, value free, uniquely determined conclusion for each case. One can quickly see why by noting that making sense of closeness to the truth really is a special case of the problem of making sense of similarity. What would it mean to say that one claim or collection of claims is closer to the truth than another? Simply that a world as described by the first set of claims is more similar than a world as described by the second set of claims to the exact way the world is. But there is no absolute notion of similarity. Any two things, or systems, are similar in countless ways, dissimilar in countless others. So any intelligible notion of "closeness to the truth" must be, at best, relative to characteristics of interest or importance. And it is doubtful that some preferred system of characteristics could be justified.[23]

There is another important respect in which there must fail to be any relevant notion of closeness to the truth. Even if we somehow settle on aspects of interest, the system that in principle delivers more accuracy in those aspects will in practice often deliver less of what is of interest in a humanly accessible way. There is a natural sense in which both special and general relativity are "more accurate" than Newtonian mechanics, in respects in which we are interested. But in most cases those very respects are more accessible to us through the in principle "less accurate" Newtonian mechanics. Thus, when human limitations are taken into account, the in-principle less accurate theory often delivers more by way of accessible approximations of human interest to "the truth" than the in-principle more accurate account.[24]

What are we to make of the fact that, at least at the level of fundamental physics, we take ourselves currently not to have any exact truths and that the idea of closeness to the truth makes sense only in a way relativised to respects of comparison, so that any evaluation of which approximation is "better" inevitably rests on application of values? Such considerations suddenly become natural instead of problematic as soon as we take on board what writers such as Cartwright and Giere have now been urging for quite some time: Science isn't in the exact truth business. It is in the imprecise and local model building business. If what we are comparing when we compare different paradigms are not candidates for truth but local and never completely precise models, valued for a wide range of both practical and cognitive features, then once the respects that we value have been specified, comparisons can be made: This model is better in THIS respect, that one in THAT respect. But the choice of respects involves application of values. Once this is recognized, a great deal of what kuhnians have urged about inter-paradigmatic theory comparisons falls into place with no threat of irrationality. Thinking of paradigms as providing incomplete and not completely exact models allows us to sidestep the problems arising from thinking in terms of truth. The problems either are avoided altogether, or, in many respects are transformed into or replaced by local, tractable, real problems that can be sensibly studied on a case-by-case basis.

The move to models is not an ad hoc resolution. It is independently motivated. Examination of how science operates shows that limited, inexact models, not exact truths, are in fact the product of the real science we have. Thinking about what science

23 See Teller (2001, pp. 402–406) for a more detailed presentation of this argument.

24 Forster (2000, pp. 241 ff.), Teller (2004, pp. 439–440).

might produce in the long run might function for some as a "regulative ideal"; but insofar as we are concerned with understanding and evaluating the structure of the science we have to hand today, speculation about what might happen in the limit of inquiry has no application.

Let me illustrate the strategy that I am advocating with our example of gravity as treated by general relativity and by Newtonian mechanics. Instead of competing truths, or candidates for truth, we now think of these approaches as providing alternative models that correspond to the world incompletely and inexactly, each in its own way. We understand gravitational phenomena better by using both kinds of models. I urge resisting the tradition according to which the account given by general relativity is somehow more accurate or more nearly correct and that there "really are no forces, instead there is curved space-time."

I have already given the general reasons against such a view, but we will better appreciate the force of these reasons by spelling them out in terms of the present specific application. General relativity is more accurate for purposes such as calculating the orbit of the planets, and quite possibly, more accurate in principle for all applications of substantive human interest. But for almost all applications we achieve accessible human accuracy only through the "less accurate" Newtonian approach. There are considerations that we clearly understand more clearly, or only, from the point of view of the general theory of relativity. For example, why can we not "shield" against gravitational forces, that is, counteract gravitational forces by forces of the same kind, as one can for electric charges or electro-magnetic radiation? We can understand this in terms of the way that massive objects "warp", or at least are correlated with the warping of space-time. A Newtonian perspective offers nothing comparable. On the other hand, consider an explanation of the tides. Few if any of us would find an account that worked entirely in terms of curved space-time to be accessible. The Newtonian account allows us to grasp the "mechanism" through which terrestrial tidal phenomena arise, and if imperfectly, still better than not at all. The Newtonian account also facilitates, in conjunction with general relativity, an understanding of why gravitational phenomena seem so force like – why they are treated so naturally by a central, inverse square law, in analogy to electrostatic forces. Altogether, piecing together a more humanly accessible and more nearly complete picture with the tools of both Newton and general relativity provides us with a much more thorough understanding than we would have if we used either of them alone.

More broadly, when we take the modelers' perspective there is no issue of potentially competing exact truths because we no longer see paradigms as providing exact truths. Instead such issues are supplanted by questions about comparing models – these are the successor real problems that I suggested are tractable, substantive, interesting, and local in the sense of calling for treatment on a case-by-case basis.

The resistance to, the horror of, considering alternative paradigms arises from the presumption that the current paradigm is "correct", in the sense of being the only approach that will yield truths about the phenomena. Such resistance dissolves when we understand the objective as instructive models instead of exact truths. With instructive models as the objective it is natural and rational to entertain alternative model building methods, if not always, at least as soon as current methods fail to produce models that address current interests and, equally, when our interests change or expand. In particular, when a paradigm produces anomalies it is defective by its own lights. In face of such defects one would be irrational not to consider alternatives when these are proposed.

Kuhnians insist that practitioners within a paradigm may sensibly prefer no solution for anomalies to a proposed solution that fails to meet their standards of explanation and other epistemic virtues. Such an insistence makes no sense insofar as the objective in view is truth. But when the objective in view is informative but inexact models, the fact that such epistemic virtues are matters of value judgment, personal or collective choice, and not rationally mandated, ceases to have the problematic consequences engendered by the objective of truth. The example of Newton's explanation of the tides again helps us see why. If competing Newtonian and relativistic accounts of gravitation are competing for the epithet, "true", we think that there has to be one right answer. The kuhnian tradition brings out the fact that each approach has things to offer that we value, but by explicitly or implicitly retaining the rhetoric of "truth" the tradition then can't make sense of the fact that the things on offer are a matter of personal or collective preference. Reasons for exact and so context and interest independent truth ought to be independent of human value considerations. Consequently, when we think of paradigms as purveyors of truths it appears impossible to reconcile paradigm choice with the fact that such choices are value laden. When we appreciate that competing – better, alternative – models have a spectrum of characteristics, some of which we value positively, some negatively, and where the characteristics as well as the balance between them can only be a matter of personal or collective preference, the application of value-laden epistemic virtues ceases to appear irrational.

Let me elaborate this diagnosis with another example. The eighteenth and nineteenth centuries saw development of extremum formulations of the Newtonian framework according to which the theory picks out the correct trajectory as the one that minimizes or maximizes some quantity. To some such extremum formulations have often seemed puzzling, even problematic. They smack of the rejected Aristotelian and Scholastic appeal to final ends, a kind of approach that the mechanical philosophy was supposed definitively to have put to rest. And some – I am one of them – have a hard time seeing such accounts as explanatory in any way. Others find them extremely illuminating. Who am I to dictate that such accounts "make no sense"?

The extremum methods are also justified by their fecundity. Some may urge that explanatory approaches can be evaluated, positively and negatively, by the extent to which they underwrite the production of models that we find informative. But no such argument is likely to pick out the uniquely "correct" modes of explanation. Witness, for example, how Heisenberg and Schrödinger's conflicting explanatory attitudes lead to the same theory!

We are tempted to think that the explosive success of the scientific revolution provides a powerful reason to think that the explanatory methods of the mechanical philosophy are right – that is truth conducive – while those of the competing Aristotelian and Scholastic tradition were bankrupt. Such a conclusion is hasty, as the example of extremum methods already shows. The history is much more complicated than the foregoing simple summary suggests. The success of Newton with planetary and a very few terrestrial phenomena aside, the successes of the scientific revolution did not really begin to add up until late in the eighteenth century, and not until the nineteenth did a veritable explosion of scientific knowledge, the "second scientific revolution", really occur. Moreover, Newton's theory is a problematic exemplar of the mechanical philosophy. Remember that his critics railed against action at a distance – precisely because it could not be seen in terms of an underlying mechanical model. Newton himself understood this all too clearly, insisted that he offered no hypotheses, that is no hypotheses about any underlying mechanical model, and advanced his account as mathematical only. He devoted most of his time to alchemy, looking for substantive explanations, be they mechanical or more in the spirit of the rejected Aristotelian holistic attitude toward natural phenomena.

Why was the mechanical philosophy, finally, so productive? Skeptics might urge that it is very hard to disentangle the role of the mechanical attitude from a host of other considerations and, should the mechanical philosophy not have dominated scientific thought, by the nineteenth century Western Europe might well have seen a different kind of flowering natural knowledge of equal practical power. Short of some more concrete scenario of how that might have happened, I find such claims far fetched, so I accept the conclusion that the explanatory precepts of the mechanical philosophy provided a powerful tool for discovery. But that is a far cry from concluding that it is the ONLY proper, correct, or epistemically productive approach to explanation, as the example of extremum methods in mechanics clearly shows.

History suggests that there are many useful ways, both instrumentally and cognitively, to build understanding of the world. The competing ways we already have to hand offer, each of them, both things we find valuable and things we regard as defects. It is a further value judgment how to balance these pluses and minuses. If the objective is, not context and interest independent exact truths, but a spectrum of inexact models that speak to human practical and cognitive interests in a large number of ways, the value laden epistemic virtues are all that one could ask of rationality.

In sum, the argument is not that the kinds of standards that we expected can, after all, be applied. These standards pertained to evaluation for truth. The epistemic virtues have always been a puzzle when truth was the objective in view. Rather, when we realize that what is to be evaluated are models that are never completely accurate in any respect and that offer a wide range of potentially desirable and problematic features, we see a substantially different roster of standards as relevant, standards that are in many respects appropriate objects of personal or community preference or that can only be evaluated instrumentally in terms of how conducive they objectively are to ends that are set by individual or collective values. I and many prior commentators have argued that there is a great deal of trans-paradigm carryover of the kinds of conclusions that function as the factual component of our evaluations. Likewise, there is a great deal of trans-paradigm agreement about instrumental claims of the form, if you want such and such kind of accommodation of such and such a phenomenon, then this model does better than that, where the kinds of accommodation that may be in question are not only numerical or more generally predictive but also explanatory. Through modelers' eyes we see that a great deal of what counts as rationality in science is instrumental rationality, that the choices we make result from an interplay between such instrumental rationality and the values imported by individuals and groups, and that such a process provides everything one could want from rationality.[25, 26]

6. PARADIGM COMPARISON THROUGH MODELERS' EYES AS A HISTORICAL AND AS A CONCEPTUAL CLAIM

The foregoing account fares poorly as a historical thesis. Advocates of the mechanical philosophy seemed to think that their approach was RIGHT, that of the Scholastics incomprehensible. Many, perhaps most practitioners in the nineteenth century thought that the Newtonian framework in particular, and the results of science very generally, provided absolute and exact truths.[27] For example, Whewell: "Now there do exist among us doctrines of solid and acknowledged certainty, and truths of which the discovery has been received with universal applause. These constitute what we commonly term *Sciences*" (In Kochelmans, 1968, p. 51). Some did see special relativity as incomprehensible, for example, William F. Magie in his presidential address to the American Physical Society in 1911: "In my opinion the abandonment of the hypothesis of an ether at the present time is a great and serious retrograde step in … physics.… How are [the supporters of the theory of relativity] going to explain the plain facts of optics?" (quoted in Van Fraassen, 2000, p. 68). It is possible that the intellectual shocks of the advent of relativistic and quantum mechanics have left important segments of our scientific community more open to alternatives than their predecessors. One does see systematic differences in attitude: Many doing work on "foundational physics" continue to think of their enterprise as getting their subject exactly right, while specialists in condensed matter physics often embrace the modelers' attitude as a common place.

But for the most part historical actors saw science as being in the truth business. The present account will be no help in understanding how they saw the methodology of their subject.

History aside, for many interpreters kuhnianism raises the question of whether major scientific change HAS to be, in some sense, an a-rational break with the past. Viewing science as a model building enterprise would appear strongly to underwrite a negative answer. The standards for what are regarded as exact truths and for what are regarded as idealized models are importantly different. Idealized models are incomplete and never completely accurate. So there is always automatically room for alternatives. Idealized models are evaluated relative to context dependent, value laden, and non-absolute

[25] Contemporary thinking about the instrumental nature of rationality in science goes back, at least, to Reichenbach's "vindication" of induction and Carnap's distinction between internal and external questions about a framework. Among many recent discussions see Giere (1989), Forster (2000, pp. 246–247), and Friedman (2001, pp. 53–55).

[26] In (2000) Bird argues that the pressure on Kuhn towards various forms of relativism can be relieved by taking a naturalistic and externalist approach to epistemology. See, especially, Chap. 6. I have considerable sympathy for that approach, but would urge that it could be further strengthened by applying the position developed here of thinking of our objective as the evaluation of imperfect models for a range of virtues, including agreement with an independent world, rather than the traditional notion of truth. In any case, I see our approaches as complementary rather than conflicting.

[27] Goethe, for example, would appear to be an exception.

standards: What are you interested in addressing? To what degree of accuracy? How do you evaluate the relative importance of the features successfully addressed? What mode of explanation functions constructively in your intellectual framework? The fact that such choices are based on values in no way compromises their rationality.

7. THE UNDERLYING PROBLEM[28]

I expect skepticism from most who are not already sympathetic to the kind of view I have been describing. I submit that the underlying basis of this skepticism turns on the received presupposition that representing and knowing about the world can only proceed in terms of unqualified truths. Knowledge requires truths, not just useful falsehoods. Explanation and understanding likewise. What we are being offered in terms of idealized models does not seem satisfactory. We want to know what things there really are in the world and what their real characteristics are. Suppose I offer you, not true statements, but false simulacra, toy models that behave only in some respects like the genuine article, and even then not exactly. This is fictionalism, a budget of phantasms; not the real thing.

Or I offer you statements that we can count as true without qualification but only, it emerges, because, sibyl like, they coyly evade any exact understanding. They are vague, inexact, not admitting of any precise interpretation. As such, I suggest, we feel that they do no better than their fictional representational alter egos. If, or to the extent that, I am right about the representational duality that I described in Sect. 2, at least for cases from science, accounts framed in imprecise terms must be subject to the same reservations that we apply to idealized models. In so far as our representations are vague and inexact, they fall short of telling us what things "really are".

Nonetheless, I submit, representation in terms of the idealized/vague IS what it is to learn about the real world, in and out of science. It is time that we stop pretending otherwise and accommodate our understanding to this circumstance.

This is strong medicine for which there is no sugar coating, but let me make an effort to facilitate the process by examining at least one kind of thinking that I take to underlie the resistance, with a view to showing where such thinking comes into question.

Let me take simple commonplaces as paradigmatic of what it is to have knowledge about the real world, commonplaces such as the fact that this piece of paper is red (think of me as holding up a large piece of bright red paper). Our criteria for knowledge of the real world should meet the standards of such examples – if this isn't knowing about the real, external, objective world, what is?

Most will find the example of a clearly red piece of paper seen in good light as an unproblematic example of knowledge. But it is not. Even without the sophisticated details, we all know that color perception is a complex matter involving interaction between objects external to us and the response of our perceptual system relative to environmental

[28] I believe that the subject of this section corresponds closely to the issue that Giere describes in terms of "perspectival realism" (1999, pp. 79–83 and to 2006).

conditions.[29] Already the mechanical philosophers understood that the so-called secondary qualities are not characteristics that can be simply attributed to external objects. So they proposed to restrict the accounts of the mechanical philosophy to descriptions in terms of the primary qualities – description in terms of shaped matter in motion.

But are there any primary qualities, in the intended sense, that is qualities that we can unproblematically attribute to external objects, independently of our conceptual apparatus? The kuhnian and kantian perspectives reject any such notion. All conceptual access is access as conceptualized. We always bring the complex world under concepts that accommodate its complexity to our sparse human limitations. Consider, for example, ordinary physical objects. We conceptualize the world in terms of discrete, determinate physical objects as in my example of the red piece of paper. But even before descending into the mysteries of quantum mechanics, any attempt to encompass such description in terms of exact truths succumbs to the problems of indefinite spatial and temporal boundaries and the puzzles of constitution. Instead, conceptualizing the world in terms of discrete, external physical objects is as much an idealization as thinking of objects as having objective, mind independent color characteristics.[30]

In both these ways, using the categories of color and of physical objects, we deal with the world in terms of idealizations. We objectify, that is we project onto those aspects of the world-as-we-know-it the idealizations as objective and exact truths. For an enormously wide range of practical and intellectual needs such objectification serves us marvelously. However, no such account is exactly correct. As we begin to take such considerations to heart, we can move towards a revised conceptualization, or model, if you will, of human representation and knowledge.

The reasoning that would appear to stand in the way of any such shift begins with the idea that exact truth is a requirement on knowledge and that our ordinary beliefs about ordinary objects serve as exemplars, as a kind of gold standard of what it is to know the world. We often paper over the conceptually shameful fact that we have no exact truths by recharacterizing what we do have – false, idealized models – as vague statements, enabling us to use the epithet "true" without qualification by hiding the lack of exact fit with a presumed independent reality under the frock of loose, not completely determinate, vague readings of our terms, reassuring ourselves with the thought that someday it shall all be made completely precise, including a precise semantics for vagueness.

Idealized models are taken to be epistemically deficient. We have believed that they do not meet the gold standard of our ordinary beliefs about ordinary objects. Hence, we have all assumed, idealized models do not qualify for more than instrumental utility. Commonplace truths fare no better. They can be taken to be true only in virtue of the slip that they get from their vagueness or semantic indeterminacy. This follows from the principle of semantic alter egos of Sect. 2 according to which the vague commonplace truths do exactly the same representational work as the exact, false idealizations to which they correspond.

[29] Giere in (2006) develops the case of color perception in detail and uses it as a vehicle to show how representation in terms of inexact models functions to provide substantive and objective knowledge of the real world, in terms of what he characterizes as "perspectival realism".

[30] Many would reject the suggestion that idealization is involved in characterizing a world as populated by determinate physical objects by appealing to one or another "4D" analysis of physical objects. I reject such accounts for reasons to be given in Teller (in preparation, b).

We must reevaluate when we realize that ordinary beliefs about ordinary objects are themselves, in fact or in effect, idealizations. Traditionally the response to such considerations has been to conclude that if even ordinary cases of knowledge do not meet the requirement of exact truth, if what we had held up as our gold standard turns out to be a kind of fools gold, then there is no knowledge, and we descend into one or another form of skepticism.

But nothing rationally mandates rejecting our gold standard when we realize that it is not quite what we thought it was. The gold standard was thought to contrast with idealizations. Then, when we discover that our gold standard – ordinary beliefs about ordinary objects – also explicitly or tacitly involves idealization, we have a choice. Retain the requirement that we thought many ordinary beliefs meet, the requirement of exact truth, and conclude that if, after all, our ordinary beliefs about ordinary objects fail this requirement, knowledge about the world is not possible. Or, retain our exemplars of knowledge of the world – ordinary beliefs about ordinary objects – as our gold standard and, realizing that they too are in fact or in effect idealizations, conclude that being an idealized model does not in itself debar a representation from counting as knowledge about the world. The world which we take direct realism to be about is already, in part, a creature of our limited and always refinable cognitive scheme. So elaborating on the commonplace scheme with further modeling need not require us to take such further elaborations to be any the less an account of what is real.

8. OF COURSE IDEALIZATIONS ARE INCOMMENSURABLE!

Incommensurability has functioned as something of a wild card in the kuhnian tradition. Interpreters, and no doubt Kuhn himself, struggled with how to understand it, what role we should take it to play in the still confused picture of trans-paradigm comparisons and paradigm choice. The considerations that I have been developing suggest a way to approach the idea of incommensurability that I believe differs substantially from prior efforts.

Start with the idea that we know the world through our never exactly correct idealizations. How can alternative idealizations be compared, given that they all fall short of exact truth and that there is no sense to be made of a context independent notion of "closeness to the truth"? We can and do make comparisons in terms of the different ways in which our idealizations contribute to and detract from a wide range of our practical and cognitive epistemic ends. These ends, in turn, constitute or rest on values that we must choose, as individuals or as groups. Truth itself, after all, was a valued end, thought to trump all others. When we appreciate that, at least today, exact truth is not humanly attainable we must choose among many other ends, ones that often compete with one another and that bear differentially on our practical concerns and our intellectual objectives. Insofar there is no "common measure", no context independent standard against which one can provide a context and value independent measure of evaluation. In this sense of "no common measure" idealizations are naturally incommensurable.

The suggestion is that the idea of incommensurability may have been driven, to a smaller or larger extent, by an implicit appreciation of the role of idealizations in our epistemic dealings with the world. But this appreciation was obscured by retaining the idea that our representations were to be evaluated for TRUTH. Once we have loosened the grip of truth as the imagined common epistemic coin we are freed to appreciate more clearly the role of complementary idealizations, the role of choice and values in their evaluation, and the eminent rationality of both.

BIBLIOGRAPHY

Bird, A. (2000) *Thomas Kuhn*. Princeton, NJ: Princeton University Press.

Cartwright, N. (1983) *How the Laws of Physics Lie*. Oxford: Clarenden.

Earman, J. and Friedman, M. (1973) The Meaning and Status of Newton's Law of Inertia and Nature of Gravitational Forces. *Philosophy of Science*, 40, 329–359.

Field, H. (1973) Theory Change and the Indeterminacy of Reference. *Journal of Philosophy*, 64, 462–481.

Forster, M. (2000) Hard Problems in the Philosophy of Science: Idealization and Commensurability. In R. Nola and H. Sankey (eds.) *After Popper, Kuhn & Feyerabend: Issues in Theories of Scientific Method, Australasian Studies in History and Philosophy of Science*. Dordrecht, The Netherlands: Kluwer, pp. 231–250.

Friedman, M. (2001) *Dynamics of Reason*. Stanford: CSLI.

Giere, R. (1989) Scientific Rationality as Instrumental Rationality. *Studies in History and Philosophy of Science*, 20, 377–384.

Giere, R. (1999) *Science Without Laws*. Chicago, IL: University of Chicago Press.

Giere, R. N. (2006) *Scientific Perspectivism*. Chicago, IL: University of Chicago Press.

Hacking, I. (1993) Working in a New World. In Paul Horwich (ed.) *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, MA: MIT. pp. 275–310.

Hoyningen-Huene, P. (1993) *Reconstructing Scientific Revolutions: Thomas Kuhn's Philosophy of Science*. Chicago, IL: University of Chicago Press.

Kochelmans, J. J. (ed.) (1968) *Philosophy of Science: The Historical Background*. New York: Free Press.

Kuhn, T. (1957) *The Copernican Revolution: Planetary Astronomy and the Development of Western Thought*. Cambridge, MA: Harvard University Press.

Kuhn, T. (1970) Reflections on My Critics. In Imre Lakatos and Alan Musgrave (eds.) *Criticism and the Growth of Knolwedge*. Cambridge: Cambridge University Press, pp. 231–278.

Kuhn, T. (1977) Objectivity, Value Judgment, and Theory Choice. In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago, IL: University of Chicago Press, pp. 321–339.

Kuhn, T. (1996) *The Structure of Scientific Revolutions, Third Edition*. Chicago, IL: University of Chicago Press.

Kuhn, T. (1993) Afterwords. In Paul Horwich (ed.) *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, MA: MIT, pp. 311–341.

Kuhn, T. (2000) *The Road Since Structure*. Chicago, IL: University of Chicago Press.

Lipton, P. (2001) Kant on Wheels. *London Review of Books*, July 19, 2001, pp. 30–31.

McMullin, E. (1983) Values in Science. In Peter Asquith and Tom Nickles (eds.) *PSA 1982*. E. Lansing, MI: Philosophy of Science Association.

Teller, P. (2001) Twilight of the Perfect Model Model. *Erkenntnis*, 55, 393–415.

Teller, P. (2004) How We Dapple the World. *Philosophy of Science*, 71, 425–447.

Teller, P. (in preparation, a) De-idealizing Truth.

Teller, P. (in preparation, b) Thinking of Objects.

Van Fraassen, B. (2000) *The Empirical Stance*. New Haven, CT: Yale University Press.

INCOMMENSURABILITY FROM A MODELLING PERSPECTIVE

Commentary on "Of Course Idealizations are Incommensurable", by Paul Teller

RONALD N. GIERE

1. INTRODUCTION

The first thing to realize about Teller's paper is that the title is somewhat ironic. He in fact argues that, given a proper understanding of scientific practice, incommensurability, understood as "semantic incommensurability", is no problem. But there is another notion of incommensurability, which Teller and some others call "value incommensurabity". In this sense, he argues, incommensurability is ubiquitous, but also not a problem. Thus, incommensurability is not a problem for a sound understanding of scientific practice.

2. SCIENCE AS MODELING

For the last 25 years, some philosophers of science, for example, Cartwright (1983, 1989, 1999), Giere (1988, 1999, 2006a), Morgan and Morrison (1999), Suppe (1989), van Fraassen (1980, 1989), and Teller himself (2001), have argued that the theoretical part of science should be thought of as involving primarily the construction and application of models. Teller provides a very good account of the modeling point of view, so I need not myself summarize this view here.

3. THE UBIQUITY OF VALUE INCOMMENSURABILITY

The most important characteristic of models, for Teller's purposes, is that they capture only aspects of parts of the world, and these never with complete accuracy. Thus, the most any scientist can conclude is that a given model is similar to some specified aspects of the world to some degree. Which aspects of the world are modeled, and to what degree

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 265–269.

of accuracy, depends in part on the individual or collective interests of scientists. And these interests in turn depend in part on the goals of particular inquiries. Since goals and interests depend on the particular context in question, there is no universal standard for saying when a model is similar enough to the subject matter under investigation. So, one claim that a particular model is similar enough is not directly comparable with a claim about a different model that it is similar enough to its subject matter. In this sense, all claims about the similarity of models to some aspect of the world are "incommensurable." Since models are by nature "idealizations," the claim in Teller's title is vindicated.

This type of incommensurability might be labeled "value incommensurability." It is similar to the incommensurability, now often called "methodological incommensurability", Kuhn (1977) saw as resulting from the application of various scientific values such as accuracy, scope, or simplicity. Kuhn argued that scientists might agree on these values but disagree on how to weigh them and trade off one against the other. Value incommensurability arises because there is no general standard as to how such weighing should be done.[1] The difference is that Teller seems to allow a much greater range of values and goals to enter into judgments of the fit of models to the world, not all of which would usually be called "scientific." For example, although a general relativistic treatment of tides on the Earth might in principle produce better fitting models, Teller claims that scientists can reasonably choose to apply less well-fitting Newtonian models because they are computationally more tractable and provide greater intuitive understanding. Such considerations seem more practical than strictly scientific.

But is this form of incommensurability acceptable? That, Teller thinks, depends on what one takes to be the aim of scientific inquiry. If the aim is the construction of well fitting models depending on various purposes, there seems to be no problem. If, however, the aim is the discovery of something like the ultimate laws of nature, or exact truths about nature, value incommensurability seems objectionably relativistic. A good part of Teller's paper is therefore devoted to examining the view that science aims at truth.

4. WHAT ABOUT TRUTH?

Teller argues that there are two complementary ways of understanding claims of truth for both ordinary and scientific statements. One way, scientific claims turn out to be strictly speaking false applied to the world, but close enough to be useful. At the same time, they may be regarded as strictly true, but only of an idealized model which fits the world only more or less well. The other way, one may take the claim to be true of the world, but then the claim itself must be regarded as vague in the relevant respects. Both ways of understanding scientific claims, Teller argues, do the same representational work. They are "representational duals" or "semantic alter egos." On this understanding of truth, there is no difference between saying that the aim of science is the discovery of truths and saying that the aim is finding well-fitting idealized models. The concept of exact truth itself is an idealized model.

[1] Wishing to avoid questions about the correct interpretation of Kuhn's own views, Teller refers to "kuhnian" themes with a small "k". He does the same for Kant. Some might object, but this practice seems justified by the diversity of interpretations and the large body of literature clearly inspired by Kuhn's work.

Teller is well aware that there is resistance among philosophers of science to this way of thinking and he explores several possible sources of this resistance. One source he does not mention is the long association of the philosophy of science with logic and the philosophy of mathematics. The concept of exact truth applies unproblematically to the abstract entities of logic and mathematics. Since the iconic theories of physics are expressed mathematically, it has been assumed that the concept of truth carries over as well. This assumption has been reinforced by the doctrine that theories are sets of statements, like axioms in a formal axiomatic system. Except for problems of inductive inference, and questions about the reality of theoretical entities, the application of theoretical statements to the world has been taken for granted. Only when some philosophers of science and others began to take seriously the practice of applying scientific concepts to the messy real world has the unreflective application of the concept of truth to scientific claims become problematic. Concern with modeling practices has been one response to this realization.

5. INCOMMENSURABILITY AGAIN

Writers on incommensurability now generally distinguish between "semantic" incommensurability, differences in the meaning of terms used in different paradigms, and "methodological" incommensurability, differences in evaluative standards in different paradigms. Teller does not explicitly invoke this distinction, concentrating instead on the more generic notion of "no common measure." He argues that it is a mistake to think that, to the extent it exists at all, such generic incommensurability poses any problems for our understanding of science.

His basic view is that there are many different sorts of relationships among theories, or paradigms, exhibiting differing degrees of incommensurability. He examines these from the stand point of modeling. At one extreme he cites the difference between the Aristotelian-Scholastic tradition and the mechanical philosophy of the seventeenth century as being almost totally disjoint. At the other extreme he considers the difference between Newtonian particle mechanics and continuum mechanics as being inconsequential. The differences between classical physics and relativistic or quantum physics are placed somewhere in between these two extremes. But he sees far less incommensurability here than is usually claimed in the kuhnian tradition. In fact he regards quantum physics and relativity physics as "expansions" or "generalizations" of, rather than a replacement for, classical physics. He sees a similar difference between extremum formulations of the Newtonian framework and standard formulations in terms of forces.

Teller traces the tendency of kuhnians to invoke replacement rather than expansion of paradigms to the fundamental idea that it is truth we are after. If, after Einstein, Newtonian mechanics is shown to be false, should it not be rejected as an account of the workings of nature and replaced with relativistic mechanics? I agree this is a factor, but I think there has been another factor at work, namely, ideas from the philosophy of language. When Kuhn first wrote, the standard view of meaning was that the meaning of any individual term is a function of its connections with all other terms in the language. So, if you change some of these connections, as one does in the move from classical to relativistic mechanics, you change the meanings of all terms in the new theory. This is a version of what we now call "semantic incommensurability." It implies replacement rather than continuity. Teller, of course, rejects strong interpretations of semantic incommensurability. He emphasizes vagueness in all claims and limiting relationships among concepts such as Newtonian mass and relativistic rest mass. In this I think he is closer to Kuhn himself than many of Kuhn's philosophical critics and expositors. Kuhn typically qualified his remarks on meaning with phrases like "to some extent," and "partially". But few people know how to understand meanings that are similar but not exactly the same.

6. THEORETICAL PLURALISM

Although he does not use these words, Teller espouses a version of what some of us call "scientific pluralism."[2] At a minimum, scientific pluralism is the view that there need not be one true theory of anything. Science can proceed with multiple theories of the same domain.

Teller considers the important example of Newtonian gravitational theory and general relativity. He rejects the commonly held view that we now know that there are no such things as gravitational forces, only curvatures in the structure of space-time itself. He argues that our understanding of gravitational phenomena is greater if we use both theories. Some phenomena, such as the bending of light near large masses, are better understood using general relativity. But other phenomena, such as the tides, are better understood from a Newtonian perspective. Our overall understanding is better if we keep both theories.

Part of Teller's insistence on pluralism is based on his rejection of the view that Newtonian gravitational theory is false and that general relativity is true, or at least closer to the truth. It is possible, however, to recover the widespread idea that general relativity is superior to Newtonian gravitational theory without invoking the notion of truth. One need only appeal to the kinds of asymmetries that Teller himself recognizes. From the perspective of general relativity we can understand why Newtonian theory works as well as it does in many applications. We can also understand things like why it is impossible to shield gravitational forces. The reverse understanding is not forthcoming. The fact, if it is a fact, that we cannot describe tidal behavior in terms of general relativity is not a defect in the theory itself but only in our limited human abilities to construct or comprehend a general relativistic account. So, given our human limitations, we have no choice but to be theoretical pluralists in approaching phenomena we want to control or understand. In my own favored idiom, we need to employ multiple *perspectives* if we are to deal as successfully as possible with the world around us.[3]

[2] This term originated with a group at the Minnesota Center for Philosophy of Science. See Kellert et al. (2006).

[3] For me, a perspective is narrower than a disciplinary matrix but broader than an exemplar. It corresponds with what in advanced sciences are often called principles. Thus, some principles of Newtonian mechanics define a set of very abstract models which by themselves cannot be applied to anything in the material world. These constitute a Newtonian perspective. These very abstract models must be supplemented with various constraints and empirical conditions to create models that can be applied to systems in the material world. See Giere (2006a, b).

BIBLIOGRAPHY

Cartwright, N. D. (1983) *How the Laws of Physics Lie*. Oxford: Clarenden.

Cartwright, N. D. (1989) *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

Cartwright, N. D. (1999) *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Giere, R. N. (1988) *Explaining Science: A Cognitive Approach*. Chicago, IL: University of Chicago Press.

Giere, R. N. (1999) *Science Without Laws*. Chicago, IL: University of Chicago Press.

Giere, R. N. (2006a) *Scientific Perspectivism*. Chicago, IL: University of Chicago Press.

Giere, R. N. (2006b) Perspectival Pluralism, in S. Kellert, H. Longino, and C. K. Waters, eds., *Scientific Pluralism*, *Minnesota Studies in the Philosophy of Science*, Vol. XIX. University of Minnesota Press.

Kellert, S., Longino, H., and Waters, C. K. (eds.) (2006) *Scientific Pluralism*, *Minnesota Studies in the Philosophy of Science*, Vol. XIX. University of Minnesota Press.

Kuhn, T. S. (1977) Objectivity, Value Judgment, and Theory Choice. In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago, IL: University of Chicago Press, pp. 321–339.

Morgan, M. S. and M. Morrison (eds.) (1999) *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

Suppe, F. (1989) *The Semantic Conception of Theories and Scientific Realism*. Urbana, IL: University of Illinois Press.

Teller, P. (2001) Twilight of the Perfect Model Model. *Erkenntnis*, 55, 393–415.

van Fraassen, B. C. (1980) *The Scientific Image*. Oxford: Oxford University Press.

van Fraassen, B. C. (1989) *Laws and Symmetry*. Oxford: Oxford University Press.

PART 10

WHAT CAN PHILOSOPHICAL THEORIES OF SCIENTIFIC METHOD DO?

THE AIM AND STRUCTURE OF METHODOLOGICAL THEORY

MARTIN CARRIER[1]

Abstract One of the challenges Kuhn's work poses to philosophy of science concerns the insight that theory-choice and, accordingly, theory-change is governed by a more complex and subtle procedure than anticipated. In particular, this procedure is claimed to inevitably and justifiedly leave room for individual preferences so that theory-choice fails to be determined unambiguously by criteria with epistemic bearing. This methodological uncertainty can be labeled as Kuhn-underdetermination. Unlike Duhem-Quine underdetermination, it does not require empirical equivalence but rather refers to a situation in which alternative theories have their strengths and faults in different areas and in different respects so that no clear overall picture emerges. Overarching methodological theories can be construed as attempts to overcome the limits set by Kuhn underdetermination. In this perspective, theories like Lakatosianism and Bayesianism provide rules for epistemic judgments that are intended to make a clear evaluation of the credentials of rivaling scientific theories possible. The two methodological theories are supposed to serve as guidelines for methodological judgment or at least to explain with hindsight why a particular theory was picked. However, on closer scrutiny the two methodological theories founder in this task of accounting for theory choice decisions. The criteria of excellence they specify are liable to uncertainties of the same sort as the more traditional virtues they are intended to replace. The paper proposes an alternative picture: methodological theories suggest general maxims and rules that guide the confirmation process rather than provide criteria for specific theory-choice decisions. Methodological theories serve to connect and unify such maxims and rules. Traditionally, lists of methodological virtues are drawn up ad hoc. One could easily add further criteria or delete others. By contrast, methodological theories provide a coherent approach to appreciating scientific theories and comparing their explanatory achievements. And they give a rationale for why these rules rather than others deserve to be preferred.

Keywords Bayesian confirmation, cognitive virtues of theories, Copernican revolution, Kuhn, Kuhn-underdetermination, Lakatos, methodological incommensurability, methodology of scientific research programs.

[1] I am grateful to the helpful comments of an anonymous referee for this volume which contributed to improving the argument.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 273–290.

1. THE CHALLENGE OF KUHN-UNDERDETERMINATION TO METHODOLOGY

In his celebrated book on *The Aim and Structure of Physical Theory* (1906) Pierre Duhem developed the important methodological insight that logic and experience alone are insufficient for assessing the cognitive credentials or the epistemic credibility of a hypothesis. Additional criteria are needed which, consequently, can only be non-empirical in kind (Duhem, 1906, Chap. X.2, X.8, X.10). This Duhemian insight constitutes the basis for one of Thomas Kuhn's most influential arguments. This argument is concerned with the nature and significance of the non-empirical criteria that are brought to bear on appraising the trustworthiness of hypotheses or theories. Kuhn argues that the procedure for assessing theories inevitably and justifiably leaves room for individual preferences. Criteria with epistemic bearing that are neutral as to the substantive commitments involved are insufficient for determining theory-choice unambiguously.

The argument roughly proceeds as follows. In Kuhn's methodological framework, the need to compare the merits of alternative theories chiefly arises during periods of crisis. In such periods, an established paradigm is undermined by an accumulation of failures in puzzle-solving, and a contender (or more than one) is advanced with the intention to supplant the received view. Under such conditions, the exclusive consideration of the empirical record does not furnish an adequate yardstick for measuring comparative success unequivocally. On the one hand, the aging paradigm suffers from a large number of anomalies; otherwise, no crisis would have emerged in the first place. The fresh rival approach was conceived only recently and pursued for a short time. For this reason it is bound to exhibit lacunae and open issues which need to be taken into account when the relative merits are to be assessed. As a result, both competing approaches are liable to exhibit empirical deficiencies. On the other hand, both approaches are also supported by some successful explanations; otherwise they would have exited from the scene earlier and never been featured as theories confronting one another in a crisis. Thus, the empirical situation must be intricate and opaque and cannot allow a clear and unambiguous ranking. It follows from the principles of Kuhn's methodological theory that logic and experience are insufficient as a basis for theory-choice. Duhem's claim becomes a theorem of Kuhn's approach.

As in Duhem, the consequence is that non-empirical virtues need to be invoked in addition. As Kuhn points out, there is a whole array of uncontroversial, prima-facie suitable standards which look impartial and are not wedded to one of the competitors. Kuhn's list of values for appraising theories, sometimes designated as "The Big Five," include accuracy, consistency, broad scope, simplicity, and fruitfulness (Kuhn, 1977, pp. 321–322). Such virtues are transparadigmatic in kind and thus avoid the charge of circularity that emerges naturally when, for instance, the commitment to push-and-shove causation is suggested as the arbiter between Cartesian vortices and Newtonian gravitational forces, or when the methodologically dubious character of mental states is advanced as a standard for deciding between behaviorism and cognitive psychology. Transparadigmatic virtues of the sort Kuhn describes are more ecumenical than such biased standards that are heavily intertwined with the substantive commitments of one of the contenders.

In addition, such virtues feature cognitive or explanatory achievements rather than social interests or aesthetic predilections. They can be linked up with epistemic aspirations or assumed goals of science (Carrier, 1986, p. 205; Hoyningen-Huene, 1992, pp. 498–499). After all, a theory that takes account of a wide realm of phenomena in a precise fashion and coheres well with other accepted beliefs yields what we demand of scientific knowledge. Such accomplishments represent what we take scientific progress to be all about. Although non-empirical in kind, such criteria arguably have an epistemic bearing.[2]

By contrast, social values or criteria of judgment are based on the appreciation of certain social or political structures. The "strong programme in the sociology of science," as inaugurated by David Bloor (1976), regards social interests as the pivot of theory evaluation. The most widely received example is the controversy between Robert Boyle and Thomas Hobbes in the 1660s about the legitimacy and bearing of experimentally gained knowledge in contrast to deductive theoretical systems. The production of experimental knowledge relied on and favored a community of "free men, freely acting, faithfully delivering what they witnessed and sincerely believed to be the case" (Shapin and Schaffer, 1985, p. 339). In contradistinction, a deductive system was based on principles laid down by philosophical masters and thus generated and sustained in an absolutist manner (ibid.). The claim advocated by Shapin and Schaffer is that the predominant political orientation of the period and experimental science had a common form of life. They were republican in spirit, not monarchist, and this shared social factor was the reason why Boylean experimentalism triumphed over Hobbesian deductivism (Shapin and Schaffer, 1985, p. 342). The contention is that the social or political struggle drives science in a particular direction; epistemic virtues do not play a significant role.

Further, the importance of aesthetic values for theory choice is advocated by the early Kuhn, the author of *The Copernican Revolution*. Kuhn claimed in this context that Copernican astronomy superseded the geocentric account because it exhibited "geometric harmony"; the initial choice between the two was a matter of taste. The Copernican arguments exclusively appealed to the aesthetic sense of the astronomers (Kuhn, 1957, pp. 171, 180). The later Kuhn was more favorable to the import of cognitive criteria. "Simplicity," to be sure, depending on how it is understood, could pass as a purely pragmatic asset that makes a theory easy to work with. Yet the remainder of Kuhn's list can be seen as indicating epistemic achievements of the theories involved.

However, given this commitment to non-empirical, cognitive virtues, Kuhn expounds two restrictions to unambiguous theory choice. The obvious impediment is that different cognitive criteria could be brought to bear. For instance, Willard V. O. Quine and Joseph Ullian feature, among other criteria, "conservatism" (i.e., coherence with the background knowledge), generality (i.e., broad scope), and "refutability" (i.e., empirical testability) (Quine and Ullian, 1978, pp. 66–80). Peter Kosso gives a different, but related cluster of cognitive criteria of evaluation which includes, among

[2] Kosso (1992, pp. 28, 35–41). Compare Worrall (2000) for contrasting views on the status of such criteria.

others, “entrenchment” (i.e., coherence with the background knowledge), precision, and generality (Kosso, 1992, pp. 35–41). It goes without saying that different lists might induce different preferences in theory choice situations.

Kuhn’s more subtle point is that even the agreement on a collection of criteria of this sort is insufficient for singling out one of the rival approaches for acceptance. The reasons are that these criteria tend to conflict with one another when applied to particular cases and that they are too imprecise to guide theory choice unambiguously. In order to appraise the relative merits of particular rival theories, such criteria need to be weighted and rendered more precise. And there is no clear rule or algorithm for achieving this task. As a result, one of the competing theories may appear superior according to some such standards and inferior according to others. It follows that transparadigmatic, epistemically relevant criteria fail to make theory choice unambiguous. There is always room left for subjective elements. This uncertainty of judgment is labeled as *Kuhn-underdetermination* or *methodological incommensurability*. The contention is that, typically, the relative merits of rival accounts do not provide a basis for unambiguously rating one of them over the other (Kuhn, 1977, pp. 322, 324–325).

Kuhn cites the competition between Ptolemaic astronomy and Copernican heliocentrism around 1550 as a case in point. As he argues, both accounts roughly coincide with respect to accuracy and scope. But geocentric astronomy outperforms its rival as regards consistency. Consistency does not alone refer to the internal structure of a theory but extends to its coherence with other views accepted in the science of the period. The notion that the earth remains at rest at the center of the universe matches Aristotelian physics excellently. According to this approach, heavy bodies fall down to the earth because they strive toward their natural place which is at the center of the universe. This Aristotelian account of the origin and nature of the weight of heavy bodies cannot be squared easily with the assumption that the earth is revolving around the center of the universe (see Sect. 2). Copernican theory suffers from its incompatibility with the physics of the time. By contrast, Kuhn goes on to argue, Copernican theory scores better regarding simplicity. At least unless simplicity is understood in terms of the computational effort it takes to arrive at some tangible result. In computational respect, namely, the simplicity of the Ptolemaic and Copernican approaches was about at the same level. Yet the latter is simpler in the sense of providing a less cumbersome account of the gross qualitative features of planetary motion (Kuhn, 1977, pp. 322–324).

What was actually at stake here was explanatory power. Theories with great explanatory power need a minimum of independent principles to account for a broad class of phenomena in an accurate fashion. Copernican astronomy excelled in this respect as regards the qualitative features of planetary motion. For instance, heliocentric theory attributes the retrograde motion of the planets, as it occurs periodically, to their order with respect to the sun and their periods of revolution. Retrogression is observed when the earth overtakes a superior planet or is overtaken by an inferior one. The core principles of Copernican theory are suitable for explaining the entirety of the relevant properties of the phenomenon, among them correlations of the sort that retrograde motion of the superior planets only occurs when these planets stand in opposition to the sun and reach their maximum brightness. In a heliocentric setting, one realizes immediately that the earth can overtake a superior planet only if the sun is located in opposition to the planet, as seen from the earth, and that the earth comes particularly close to the planet under such circumstances. Ptolemaic astronomy was also able to cope with these correlations, to be sure, but it had to resort to an extra hypothesis for each such account. In contrast to the Copernican scheme, Ptolemy needed additional, tailor-made assumptions for every single aspect of the phenomenon. In view of achievements of this sort, Copernican astronomy clearly stands out in explanatory power (Carrier, 2001, pp. 81–92).

Anyway, the picture presented by Kuhn takes Ptolemaic and Copernican astronomy to be roughly on a par as regards accuracy, and which one of the two (if any) counts as being more simple depends on how “simplicity” is understood or rendered precise. Further, heliocentrism lags behind as to consistency, i.e., its coherence with other parts of accepted scientific theory, but displays a paramount power of explaining astronomical phenomena. There are no geocentric anomalies resolved, but the qualitative heliocentric explanations are more stringent and do not turn on ad-hoc adaptations. Judged in terms of these cognitive values, the comparative assessment leads to a draw.

The example of the Copernican Revolution is atypical in that the rival accounts entailed approximately the same observational consequences (at least with respect to astronomical phenomena). Yet unlike Duhem-Quine underdetermination, Kuhn-underdetermination does not require empirical equivalence but rather refers to a situation in which alternative theories display their strengths and weaknesses in different fields and in different respects so that no clear overall comparative assessment of the epistemic accomplishments is feasible. This comes out more distinctly in Kuhn’s second example, namely, the Chemical Revolution (Kuhn, 1977, p. 323). The phlogiston theory had a hard time accommodating the weight increase during combustion and calcination whereas the oxygen theory was at a loss to account for the fact that the properties of the supposedly elementary metals resembled one another much more closely than the properties of the corresponding oxides – in spite of the fact that the latter were assumed to contain oxygen as a common ingredient. By contrast, this trait appeared quite plausible on the phlogistic principle that metals derived their chief properties from phlogiston as their common constituent.

Accordingly, both contenders were faced with anomalies – albeit of a different sort. Yet the impact of these anomalies was far from obvious and depended on how the commitment to empirical adequacy, shared by the competitors, was spelled out. The chief aspiration of the phlogiston theory, and the yardstick used for measuring its achievement, was to explain, in a qualitative fashion, the properties of the chemical substances involved and their changes during chemical reactions. In this respect, phlogiston theory outstripped its rival – as the example suggests. Oxygen theory, by contrast, placed quantitative weight relations at the focus and in this respect greatly exceeded the performance of the traditional view. As a result, taking variations of empirical adequacy into account does not change the conclusion suggested by the Copernican Revolution. The notion of empirical adequacy may be understood differently, just as the non-empirical virtues addressed before, and is in need of adjustment if it is to provide an unambiguous balanced appraisal of the empirical accomplishments of the theories involved.

The upshot is that methodological judgment is rendered ambiguous by the room left for selecting and balancing different criteria of evaluation. The adoption of different standards may issue in divergent preferences, and even a shared canon of criteria or values for choosing between competing theories fails to provide a clear measure of epistemic achievement. How such theories are to be rated depends on how the relevant values are rendered precise and how they are weighted, i.e., which one is given precedence in case of conflict (Laudan, 1984, p. 83; Kosso, 1992, pp. 46–47; Nola and Sankey, 2000, p. 28). No unambiguous comparative evaluation emerges from bringing Kuhn's "Big Five" to bear. Duhem's claim is sharpened to the effect that logic, experience and the commitment to cognitive values are jointly insufficient to establish a clear rank-order among rival theories. The methodological judgment is Kuhn-underdetermined.

It needs to be underlined that Kuhn himself takes the room for theory choice opened up by Kuhn-underdetermination as an asset rather than a liability. The reason is that the prospects of success for rival accounts are typically objectively indeterminate. Nobody is in a position to reliably anticipate which one will prevail in the end. A procedure of judgment that yields unambiguous scores at the early stages of theory articulation is liable to mislead the scientific community into pursuing only one option – the allegedly superior one. If it turns out later that the choice was mistaken, no theoretical resources are left for correcting the error. In a situation governed by uncertainty, the optimum strategy is spreading the risk so that different parts of the community explore different theoretical tacks. This is achieved naturally if the competitors score differently in different respects. Given the usual variation in individual preferences, Kuhn underdetermination can be expected to produce the desired division of the community. Individual uncertainty leads to the collective epistemic optimum (Kuhn, 1970b, p. 248, 1977, pp. 325–334; Hoyningen-Huene, 1992, pp. 493–494; Carrier, 2002, p. 58).

Kuhn suggested that theory choice and, consequently, theory change was governed by a more complex and subtle procedure than anticipated by earlier methodological accounts. The underdetermination of judgment creates room for science to cope appropriately with epistemic risk and uncertainty. By contrast, others considered Kuhn-underdetermination a major threat to scientific rationality. Imre Lakatos claimed that only universal epistemic standards of judgment are compatible with the requirements of sound methodology. Any room left to personal predilections involves a surrender of science to authority. The judgment of the scientific community would be shaped by the opinion leaders and theory change would become a matter of "mob psychology" (Lakatos, 1970, pp. 90–91, 1973, pp. 324–325). In other words, Kuhn-underdetermination is regarded as the first step toward relativism and the collapse of scientific rationality. In view of the contentious import of Kuhn-underdetermination, it appears worthwhile to examine options for blocking its emergence.

2. METHODOLOGICAL THEORIES AND HISTORICAL THEORY CHANGE

Kuhn's argument proceeds from a collection of diverse cognitive values. Their lack of precision and the possible conflict among them is the reason why Kuhn-underdetermination or methodological incommensurability can arise in the first place. Consequently, its emergence could be blocked by introducing criteria of judgment that are sufficiently precise and of such a sort that no discordant assessments are entailed by them. Overarching methodological theories can be construed as attempts to provide such an advanced basis for the evaluation of scientific achievements and thus to overcome the limits set by Kuhn-underdetermination. In the following, I turn to Lakatos' methodology of scientific research programs and to Bayesian confirmation theory as examples of such attempts. Lakatosianism and Bayesianism can be interpreted as endeavors to make a rule-governed, epistemically based, and unambiguous assessment of the merits of rival scientific theories possible and thus to neutralize the impact of Kuhn-underdetermination. Let me explore how well the two methodologies fare in this respect.

Lakatos' unit of methodological appraisal is the research program or, more precisely, the series of theories that make up such a program. A research program is characterized by a "hard core" of principles which are retained by its advocates at all cost. Problems are taken care of by adapting the "protective belt" of auxiliary hypotheses which is taken to consist of observation theories, suppositions of initial and boundary conditions, and additional assumptions within the corresponding theory itself. In conceptual respect, Lakatos demands that the program development be directed by a "positive heuristic," i.e., a set of comprehensive theoretical guidelines for its future development. The positive heuristic singles out significant theoretical problems and offers tools for their solution. It directs scientists through the maze of confusing difficulties by providing a plan how to elaborate the program. One of Lakatos' historical assertions is that the development of a distinguished, high-quality program is determined by its positive heuristic and does not merely respond to conceptual and empirical difficulties. Methodological standards are supposed to direct the transition from a program version to its successor or the replacement of a program by a competitor. I focus on the most important standards as specified by Lakatos. An acceptable successor version within a program is required, first, to remain in agreement with the positive heuristic of the program, second, to account for all those phenomena that are successfully explained by its predecessor – albeit, possibly, in a different fashion, and, third, to successfully predict some novel, hitherto unexpected facts. Lakatos demands the reproduction of the previous empirical achievements and the anticipation of observational regularities that were unknown to science before. Rival research programs are to be evaluated analogously. Programs that satisfy these program-internal standards are compared by applying the conditions of reproduction and anticipation to the competitors. A superior program has to reproduce the explanatory successes of the rival program and predict effects unexpected within the framework of the latter (Lakatos, 1970, pp. 33–36, 47–52, 68–69; Carrier, 2002, pp. 59–63).

Kuhn grounds the comparative evaluation of theories on a collection of alternative criteria – which is why conflicts among them can arise. By contrast, Lakatos requires the joint satisfaction of a number of demands. No conflict between the reproduction of explanatory content and the anticipation of novel regularities can possibly surface because both conditions need to be met. The Kuhnian process of weighting alternative standards is avoided altogether. Assuming that Lakatos' requirements are sufficiently precise, Kuhn-underdetermination cannot become manifest (see Carrier, 2002, pp. 63–64).

Consider the Copernican program at the time of Galileo as an example. The theory was able to account for the entirety of planetary positions with roughly the same (or even slightly increased) accuracy as compared to the Ptolemaic program. In addition, as Galileo showed, the Copernican theory entailed a stunning new effect, namely, the phases of Venus. Neither this consequence of the Copernican scheme nor the existence of the phenomenon had been realized before. According to the Ptolemaic theory, Venus was assumed to revolve around the earth and at the same time perform a rotating or "epicyclic" motion in the course of this revolution. In addition, since Venus always remains in the vicinity of the sun (after all, it features as the "morning star" or the "evening star," respectively), Ptolemaic theory surmised a coupling of its motion around the earth to the position of the sun. The overall result was that Venus was taken to rotate before or below the sun. Consequently, the appearance of Venus should oscillate between a phase of darkness analogous to new moon and a crescent-shaped phase. On the heliocentric account, by contrast, Venus is supposed to revolve around the sun and should run, accordingly, through the full cycle of phases like the moon. All these consequence relations had escaped the attention of astronomers prior to Galileo, and the phenomenon itself is unobservable with the naked eye. In 1611, Galileo discovered the effect using a telescope, he realized that it followed from the Copernican scheme and presented it as a confirmed prediction and a major epistemic achievement of heliocentrism. Indeed, in view of Lakatos' criteria of reproduction and successful anticipation, Copernicanism is rightly considered superior to the Ptolemaic account.[3]

Lakatos' methodology thus appears to specify fairly unambiguous evaluation rules for competing accounts. But it does not fall victim to Kuhn's charge that any evaluation algorithm may lead to rash and inappropriate judgments. The reason is that Lakatos' approach automatically implies an initial limitation of methodological evaluation. According to Lakatos, empirical support primarily arises from confirmed predictions. However a certain period of time has to be granted for ascertaining whether a theoretical anticipation is borne out empirically. As a consequence, the performance of a theory cannot be assessed instantly. This delay in judgment Lakatos calls "the end of instant rationality" (Lakatos, 1970, pp. 68, 87). It entails that Kuhn-underdetermination may be resolved with hindsight, but continues to restrict actual theory-choice decisions.[4] On the whole, then, Lakatos' claim is that comparative theory assessments can be reconstructed as rule-guided and rational without at the same time being liable to Kuhn's charge of prematurity. Once in a while it seems possible to have one's cake and eat it, too.

[3] The situation is complicated through the fact that the geoheliocentric, "Tychonic" successor version to the Ptolemaic variant of geocentric astronomy also yielded this confirmed prediction.

[4] In addition, Lakatos wants to separate methodological evaluation from any recommendation to the relevant scientists regarding which program to accept or to pursue. No advice is given to accept or pursue the superior theory. The reason Lakatos advances for abstaining from any such advice is its fallibility that is due to the fact that apparently unsuccessful predictions may prove correct after all and that the success of a program is crucially dependent on the resources and the creativity devoted to its development (Lakatos, 1970, pp. 68–70, 1978, p. 217).

However, on closer scrutiny Kuhn-underdetermination resurges in Lakatos' methodological framework. Its chief source is the change in the scope of programs or the division of programs. Consider a comprehensive program that is followed by a number of narrower successor programs so that the domain of the former is later covered by several distinct approaches. This situation is not untypical and it tends to reintroduce the challenge of ambiguous methodological judgments. The shift toward heliocentrism is a case in point. The pertinent original program is Aristotelian physics and cosmology. One of its principles says that all bodies strive toward their natural places where they come to rest (see Sect. 1). The natural place of "heavy" bodies (in contrast to "light" ones such as air and fire) coincides with the center of the universe. It follows that the earth is located at this center and that solid or liquid (i.e., heavy) bodies move on their own in the direction of this center and, consequently, toward the surface of the earth. Given this framework, the excentric location and the revolution of the earth should produce effects on terrestrial motions which are not, in fact, observed. If the sun were at the center of the universe, heavy bodies left to themselves should move there rather than fall on the surface of the earth. And if the earth were in motion, bodies thrown perpendicularly into the air should be deviated in westward direction. After all, the trajectory toward the center of the earth remains unaffected by an eventual revolution of the earth so that its surface should proceed eastward while the body is in flight. No one of these consequences is confirmed empirically and these anomalies militate against the Copernican allegations.

It follows that if the domain of application of heliocentrism is limited to astronomical issues, its greater explanatory power makes it look superior to the geocentric alternative (see Sect. 1). But if the whole of physics and cosmology is taken into consideration, the higher degree of coherence of the geocentric approach or, conversely speaking, the occurrence of anomalies to heliocentrism in terrestrial physics suggests that geocentrism retains the lead. That is, the two theories are ranked differently according to the scope that is deemed relevant. Kuhn-underdetermination again makes its appearance.[5]

This is by no means a singular case. Consider Hendrik Lorentz' classical electron theory that was superseded in part by special relativity theory and in other respects later by quantum mechanics. Special relativity is the heir to electron theory as regards the electrodynamics of moving bodies but has nothing to say on such interactions between charges and fields that become manifest, for instance, in the normal Zeeman effect (i.e., the split of spectral lines in a magnetic field). This effect was accounted for in Lorentz' theory and later incorporated into quantum mechanics. But how are we to judge the situation around 1910 when quantum mechanics was not yet in the offing? If we proceed from the entire domain of application of electron theory, we need to conclude that special relativity failed to fully reproduce the explanatory success of the earlier account. In view of the fact that special relativity comes out superior within its domain, the upshot is that no clear rank-order among the two approaches can justifiably

[5] It was first systematically elaborated by Larry Laudan that the domain of application of a theory may be contentious and that the scope attributed to it may strongly influence its methodological appraisal (Laudan, 1977, pp. 19–21).

be specified. If, by contrast, consideration is limited to the domain of special relativity, electron theory is superseded. Again, the methodological judgment is dependent on which scope is assigned to the rival programs.

A third example is taken from psychology and concerns the "cognitive revolution" of the 1960s which involved the replacement of behaviorism with cognitive psychology. The development of cognitive psychology brought major advances in the explanation of human behavior – including the successful prediction of novel facts (Carrier and Mittelstrass, 1991, pp. 132–140; Carrier, 1998, pp. 223–224). On the other hand, the new program largely failed to extend to animal behavior, so that with respect to the latter, behaviorism essentially remained in place. It didn't appear overly plausible to attribute elaborate motivational processes and other intricate mental procedures to pigeons, rats, or the common fruit fly. With respect to these latter creatures, Skinnerian reinforcement mechanisms were retained. As a result, the comparative evaluation is crucially influenced by what is taken as the relevant domain of application. If the scope is limited to humans, cognitive psychology clearly supersedes behaviorism; if the relevant domain is extended to all living creatures, cognitive psychology fails to reproduce the whole of the behaviorist explanatory content so that no clear judgment emerges.

Speaking more generally, the scope dependence of judgment reveals an ambiguity or lack of precision in Lakatos' criteria of evaluation that might lead to a shifting ranking of theories. In particular, the requirement of explanatory reproduction can be interpreted differently. It can be restricted to narrower domains or extended to broader ones. As a result, Kuhn-underdetermination proves resistant to facile dissolution.

Problems of Bayesianism likewise emerge primarily with respect to the precision of methodological criteria – or lack thereof. Bayesianism takes Bayes' theorem of probability theory as the basis for methodological judgment. Construed along such lines, Bayes' theorem says that the probability of a hypothesis given the available evidence, *p(h/e)*, equals the likelihood of the evidence, that is, the expectedness of the data given the hypothesis, *p(e/h)*, times the prior probability of the hypothesis, i.e., its probability before the evidence *e* was actually registered, *p(h)*, over the probability of the evidence, *p(e)*:

$$p(h/e) = \frac{p(e/h)\,p(h)}{p(e)}$$

Bayesianism is committed to using Bayes' theorem as a rule for adapting and updating beliefs in light of the evidence. This so-called principle of conditionalization directs us to employ Bayes' formula for assessing assumptions. Hypotheses are the better confirmed the higher their probability is (Earman and Salmon, 1992, pp. 89–92).

In opposition to Lakatosianism, Bayesianism specifies a number of different features of confirmation which may contrast with one another. It could well be the case that a hypothesis is superior to an alternative as regards prior probability but lags behind with respect to the specification of relevant evidence. Yet Bayes' theorem outlines a mathematical procedure for balancing the various features involved.

Let me assume for the sake of argument that the challenge of weighting divergent requirements unambiguously, as highlighted by Kuhn, can be met within the Bayesian framework. What is more worrying is the second one of Kuhn's challenges, namely, to render precise the criteria used for rating hypotheses. Any Bayesian hypothesis appraisal requires the evaluation of all the quantities contained in Bayes' theorem. But real-life examples presented by Bayesians themselves (Howson and Urbach, 1989, pp. 96–102) make it conspicuous that assigning quantitative measures to the Bayesian magnitudes is a highly arbitrary affair. It is far from obvious, for instance, what is a fair estimate of the prior probability of the heliocentric hypothesis around 1550. And it is quite tricky to assess the expectedness of phenomena like retrograde planetary motion in the Copernican scheme as compared to the Ptolemaic one. A large amount of deliberate fine-tuning is requisite until a rank-order of hypothesis probabilities is finally produced. The suspicion is that different estimates can be given within the bounds of plausibility which would reverse the order of hypothesis probabilities (Worrall, 1993, pp. 333–342).

As a result, Kuhn-underdetermination is here to stay. Actually, this feature is sometimes advanced as a virtue rather than a vice of Bayesianism. For instance, Wesley Salmon argues that the prior probability in Bayes' theorem designates the plausibility of the hypothesis against the relevant background knowledge. Such estimates of plausibility are clearly liable to subjective variation. It follows that the evaluation of quantities in Bayes' theorem leaves room for individual choice. The application of the same formula to the same methodological challenge may thus yield divergent methodological assessments. This helps to avoid the risk of premature unanimity, as described by Kuhn. It is at this juncture where Tom Kuhn meets Tom Bayes (Salmon, 1990, pp. 180–182).

To sum up, it seemed at first sight that the limitations in methodological judgment that underlie Kuhn-underdetermination originated from the haphazard character of Kuhn's list of criteria. Accordingly, it appeared promising to overcome these limitations by appealing to systematic methodological theories. I concentrated on two such theories. Lakatos focuses on confirmed predictions of novel facts, Bayesianism places hypothesis probability at center stage. These theories were supposed to direct methodological judgment or at least to explain with hindsight why a particular scientific theory was adopted under the conditions at hand. However, on closer scrutiny both methodological theories founder in this task of accounting for theory choice decisions. The criteria of excellence they specify are liable to uncertainties of the same kind as Kuhn's accidental collection they were supposed to replace. Lakatos' methodology suffers from the uncertainties involved in delineating the scope of research programs, and Bayesianism comes to grief because of the arbitrariness of imputing numerical values to all conditional probabilities needed in order to arrive at hypothesis probabilities.

At a second glance, however, this null result makes sense. Kuhn's argument as to premature unanimity provides an epistemic reason why appropriate methodologies ought to fail in rank-ordering hypotheses or theories unambiguously. This suggests that methodologies should abstain from the attempt to guide theory choice stringently. Accounting for theory choice might have been the wrong job description for methodologies in the first place. Let's examine, therefore, whether methodologies prove helpful in another respect.

3. HIGHLIGHTING FEATURES OF EXCELLENCE

An alternative picture is that methodological theories highlight features that are relevant to the excellence of a scientific theory. They indicate what a scientific theory worth being accepted should be like and which credentials rightly count in its favor. Conversely, methodological theories specify which factors have no impact on the quality assessment of a scientific hypothesis. The particular asset of methodological theories is that they serve to connect and unify such features of excellence. Kuhn's collection is assembled ad hoc. One might easily add further criteria or delete others. By contrast, a methodological theory identifies such features from a unified point of view. It gives a systematic and coherent account of methodological distinction and thus provides a rationale as to why these features and not others are to be preferred. Yet when it comes to drawing on such criteria for selecting the most appropriate hypothesis or theory under particular circumstances, no algorithm or straightforward procedure can be articulated. Methodological theories fail to remove Kuhn-underdetermination, but they can contribute to designating the space left to individual judgment and choice more systematically and more thoroughly.

Let me begin with the bearing of anomalies on the appraisal of theories. Popper famously required that all empirical difficulties be taken seriously. Scientific method demands not glossing over anomalies but dealing with them and treating them as potential counterexamples. Kuhn, by contrast, advances an immunity claim of paradigms to anomalies as a historical generalization, and he adds the epistemic rationale that anomalies are ubiquitous in science so that the advice to accept each of them as a potential refutation would be tantamount to closing down the business of theory construction (Kuhn, 1970a, pp. 79–82). Here is what follows on this issue from Lakatos' methodology. A program or a program version is backed by correctly anticipated empirical regularities. Assume a contender challenges the theory at hand. Successfully predicted facts have not been accounted for by the contender; otherwise, they would not have been novel in the first place. Consequently, the methodology implies that only those facts are suitable for buttressing a program which cannot be explained by its rival. Reversing the point of view, it follows that only those facts militate against a program that favor its competitor. This entails that only those anomalies count as a failure which can be solved by the rival in an acceptable fashion, i.e., by predicting some novel property of the phenomenon or some hitherto unknown phenomenon. In sum, no refutation without confirmation. Accordingly, the mere inability to accommodate this or that observational regularity does not bring a program into trouble (Lakatos, 1970, pp. 37, 92).

On the whole, then, Kuhn's immunity claim can be derived from Lakatos' requirements for empirical support. Research programs are rightly immune to mere anomalies. If a research program is to be criticized effectively it is not sufficient to expound its liabilities. Rather, what hurts a program is the methodologically distinguished solution to its problems within a different theoretical perspective.

Another historical regularity stressed by Kuhn is that paradigms are never given up unless an appropriate alternative is available. Scientific revolutions always involve theory-substitutions. Its Lakatosian analog is: No program abandonment without program replacement. A program is only discredited methodologically if a superior competitor is waiting in the wings. This condition can be derived from a corollary to the immunity argument. This argument said that the liabilities of one theory are the assets of the other. There are no significant failures without an alternative solution. And obviously enough, if a theory is not in trouble it should not be given up. It follows that a program can never be rated as deficient unless there is a contender attacking it with some success. Disconfirmation of a program is produced by a corroborated rival.[6]

Kuhn's claims as to paradigm immunity and paradigm substitution are leveled as factual objections to methodological requirements entertained by Popper. They are advanced as historical counterexamples to Popper's demands. Within the framework of Lakatos' methodology, by contrast, the two features of immunity and substitution constitute theorems rather than objections. Lakatos' conception thus provides a methodological explanation for these Kuhnian characteristics of scientific change (Carrier, 2002, pp. 64–66).

The positive scheme Lakatos devises places emphasis on the planned growth of knowledge as a chief distinction of an acceptable research program. I mentioned that the central empirical hallmark suggested by Lakatos is the confirmed prediction of novel facts. This brings patterns of theory change to the focus of methodological assessment. Mature science contains coherent, large-scale research programs, unified by a set of principles and directives; superior programs do not respond to the data but anticipate them. This suggests that Lakatos' methodology outlines desirable features of a theory. The methodology says in a systematic and well-articulated fashion what a good theory should be like and which credentials can be rightly cited in its favor. It is true, methodologies thereby automatically contribute to explaining certain patterns of theory choice. If a theory is accepted by the scientific community which excels in these respects, the methodology serves as an explanation of this choice. In view of the argument in Sect. 2, however, it has to be granted that there will always be a number of opaque and messy cases in which the actual choice is underdetermined by the criteria suggested by the methodology. This is why the specification of the attractive features of a theory falls short of accounting for theory-choice decisions comprehensively.

The same lesson emerges with respect to Bayesianism. In order to realize the relevant features more distinctively, Bayes' theorem needs to be developed mathematically. If the denominator is expanded using the theorem of total probability, the so-called second form of Bayes' theorem is obtained which was first enunciated by Pierre Simon de Laplace (Howson and Urbach, 1989, pp. 26–27, 86–87).

$$p(h/e) = \frac{p(e/h)\,p(h)}{p(e/h)\,p(h) + p(e/\neg h)\,p(\neg h)}$$

[6] Lakatos (1970, p. 35, 1971, pp. 112–113). It follows that any serious test of a theory demands a comparison between theories. This important consequence of Lakatos' methodological approach completely escaped Kuhn's notice. Kuhn mistakenly objected that in Lakatos, as in Popper, the existence of a theoretical alternative was merely accidental (Kuhn, 1980, p. 191). In fact, Lakatos' notion of empirical support is crafted such that only a successful competitor can bring a theory into trouble.

"$\neg h$" or "not h" means that the hypothesis is wrong; $p(e/h)$ and $p(e/\neg h)$ express the degree to which the data e can be expected to occur if the hypothesis was true or false, respectively. Rearranging the terms on the right-hand side yields:

$$p(h/e) = \frac{1}{1 + \frac{p(e/\neg h)\, p(\neg h)}{p(e/h)\, p(h)}}$$

This form of Bayes' theorem makes it conspicuous that the hypothesis probability or degree of confirmation depends on two ratios. First, the ratio of the likelihood of the evidence given the falsity or the truth of the hypothesis, respectively: $p(e/\neg h): p(e/h)$. This ratio expresses to which degree one could anticipate the occurrence of the data e by relying on h or, alternatively, by drawing on the background knowledge alone. The resulting hypothesis probability is high if this ratio is small, that is, if the likelihood of the data is small if h is false but large if h is true. Second, the ratio of the prior probability of the hypothesis to its negation: $p(\neg h): p(h)$. The more plausible h was before any relevant data turned up (and, correspondingly, the less probable $\neg h$ was), the better h is confirmed ceteris paribus.

I focus on the influence of the likelihood, that is, the expectedness of the data. The crucial feature is that data which are not to be expected given the background knowledge alone, but are suggested if a hypothesis is accepted in addition, have a particularly strong confirmatory impact. It is the increase in the expectedness of the data through the adoption of the hypothesis that makes this hypothesis probable. Let me give two examples of this feature.

The increase in likelihood is the methodological procedure Bayesianism employs for granting excellence to theories which account for otherwise mysterious phenomena. Such phenomena are surprising given nothing but the pertinent background knowledge. That is, the likelihood of the relevant data is very low. Let some hypothesis be adopted that explains these data (if conjoined to the background knowledge). Given this amended knowledge, the effect was to be expected. The likelihood of the evidence is raised by accepting the hypothesis, so that the probability of the hypothesis is thereby increased.

The same procedure also covers the methodological distinction of confirmed predictions of novel facts. The adoption of a hypothesis which anticipates new observational regularities obviously heightens the expectedness of these regularities. From a Bayesian point of view, it is immaterial for this increase in likelihood whether or not the evidence was known before the hypothesis was formulated. The occurrence of the full cycle of the phases of Venus was certainly not to be expected in the Ptolemaic approach, but it could have been anticipated on a Copernican basis. Galileo's observation thus supports Copernicanism on both Lakatosian and Bayesian grounds although the mechanism of support is different in each case. In Lakatos, temporal novelty plays a key role, in Bayesianism the increase in the expectedness of data, possibly well known before, is the critical quantity.

The same feature of likelihood increase also underlies the Bayesian appreciation of theories that contribute to unifying otherwise diverse phenomena. The basis for this methodological premium awarded to theoretical unification is a characteristic of probability theory according to which the joint occurrence of independent events is less probable than their individual occurrence. The basis is the multiplication theorem of probability: $p(a \wedge b) = p(a)\ p(b)$. If there is no connection among phenomena, the likelihood of their combined appearance, $p(a \wedge b)$, is lower than that of their separate appearance, $p(a)$ or $p(b)$, respectively. By contrast, if a theory reveals that these phenomena are related to one another, their joint occurrence is anticipated. They were not independent events in the first place but are rather produced by the same process. As a result, the likelihood of the entire class of phenomena is raised, and this significant increase strongly confirms the underlying unifying hypothesis.

Actually, this type of explanatory achievement was claimed by Copernicus and the early Copernicans as the major advantage of heliocentrism (Carrier, 2001, pp. 134–136). I mentioned before that the heliocentric account of retrogression naturally established the observed correlations between the relevant aspects of the phenomenon, and that the Ptolemaic approach had to adjust these correlations by hand (see Sect. 1). For instance, it follows immediately from the heliocentric order of the planets that if the earth overtakes a superior planet, retrograde motion occurs, the sun is located opposite to the planet, and the planet is closest to the earth and thus appears most brightly. In the Ptolemaic framework, such connections among the phenomena are taken care of by specific, tailor-made adaptations. There was no reason to expect these correlations prior to their actual observation. Analogously, it follows from the same principle of the heliocentric order of the planets that Mercury and Venus always remain in the vicinity of the sun – their elongation is bounded – whereas the angular distance of the other planets is unlimited. This distinction among the planets can be reproduced on a Ptolemaic basis, to be sure, but only by adding specific assumptions. Copernicans gave a unified explanation of this set of apparently diverse phenomena. They thus increased the likelihood of these phenomena and thereby conferred a high degree of credibility to the unifying hypothesis.

4. CONCLUSION

In this essay I have approached the issue of the appropriate role of methodology from a basically empirical perspective by exploring the bearing of diverse methodologies on theory-choice in science (Carrier, 1986, pp. 210–220; Nola and Sankey, 2000, pp. 22–26). The result is, first, that methodological theories are not particularly good at giving concrete advice in such matters. The strength of methodological theories does not lie in singling out the best scientific theory unambiguously or in attributing precise degrees of confirmation. It requires art of judgment to bring general considerations to bear on particular, intricate cases. Attempts to reach a comparative assessment of rival theories by straightforward derivation from methodological principles are thwarted, as a rule, by the complexity of the situation. Meanderings in the program structure often prohibit the unambiguous assessment of methodological assets and liabilities; the multiplicity of alternative criteria and their lack of precision makes a clear, rule-governed verdict hard to reach. Methodology falls short of providing an evaluation algorithm; Kuhn underdetermination will continue to haunt methodological assessment.

However, second, methodological theories still play an important role. Their strength becomes conspicuous once we compare methodologies such as Lakatosianism and Bayesianism with Kuhn's list of explanatory virtues. Kuhn's Big Five constitute an accidental collection of methodological values that might be expanded or abridged according to day-to-day needs. They are not tied together by a unifying principle. In contradistinction, methodological theories specify a coherent set of explanatory virtues. They systematically determine which features contribute to epistemic merit.

For instance, Lakatosianism takes theory change as the key element for judging epistemic achievement (see Sect. 3). This general commitment ramifies into emphasizing the guidance of theory development by heuristic directives and the successful prediction of formerly unknown phenomena. Analogously, Bayesianism is led by its commitment to Bayes' theorem to stress virtues to the effect that a good hypothesis should be plausible in light of the background knowledge and show that the data could have been expected beforehand. Methodological theories serve to endow a set of such features of excellence with a unifying perspective that makes it clear why precisely this set is an appropriate indication of epistemic achievement.[7]

In his famous study on the *The Aim and Structure of Physical Theory* Duhem considers it the chief objective of physics to order and classify empirical generalizations. The aim and structure of methodological theory comes out not completely dissimilar. Its principal task is to give a systematic and unifying account of the diverse rules, maxims, virtues, and criteria that might be considered relevant to judging the epistemic achievement of a theory. The basis of this understanding is Duhem's insight, as mentioned before (see Sect. 1), that non-empirical virtues are inevitably to be drawn upon, in addition to logic and experience, in order to assess the credibility of claims in science. But Duhem advances only the *bon sens* of the scientists as a yardstick. Kuhn goes further in this respect in making the relevant virtues more explicit, but the collection of epistemic values he suggests remain unsystematic and lack a unifying perspective. It is the challenge of methodological theories to do better and to achieve for criteria of judgment what scientific theories, following Duhem, accomplish for empirical generalizations: establishing a systematic order and classification.

[7] The meta-methodological principle underlying this approach to evaluating methodological theories is "reflective equilibrium." This principle directs us to bring intuitions about particular instances into harmony with intuitions about general rules. Transferred to the realm of methodology, the task is to preserve a maximum of normative intuitions about particular scientific achievements and about general cognitive goals of science (Nola and Sankey, 2000, p. 25). It is in virtue of this principle that the inability to justify concrete instances of theory choice in an unambiguous way, as elaborated in this essay, creates a tension which is attempted to be resolved by a changed assessment of the proper role of methodologies.

BIBLIOGRAPHY

Bloor, D. (1976) *Knowledge and Social Imagery*. London: Routledge.

Carrier, M. (1986) Wissenschaftsgeschichte, rationale Rekonstruktion und die Begründung von Methodologien, *Zeitschrift für allgemeine Wissenschaftstheorie*, 17, 201–228.

Carrier, M. (1998) In Defense of Psychological Laws. *International Studies in the Philosophy of Science*, 12, 217–232.

Carrier, M. (2001) *Nikolaus Kopernikus*. München: Beck.

Carrier, M. (2002) Explaining Scientific Progress. Lakatos's Methodological Account of Kuhnian Patterns of Theory Change. In G. Kampis, L. Kvasz and M. Stöltzner (eds.) *Appraising Lakatos: Mathematics, Methodology, and the Man (Vienna Circle Library)*. Dordrecht, The Netherlands: Kluwer, pp. 53–71.

Carrier, M. and Mittelstrass, J. (1991) *Mind, Brain, Behavior. The Mind-Body Problem and the Philosophy of Psychology*. New York: de Gruyter.

Duhem, P. (1906) *The Aim and Structure of Physical Theory*. New York: Atheneum, 1974.

Earman, J. and Salmon, W. C. (1992) The Confirmation of Scientific Hypotheses. In H. S. Merrilee (ed.) *Introduction to the Philosophy of Science*. Englewood Cliffs, NJ: Prentice-Hall, pp. 42–103.

Howson, C. and Urbach, P. (1989) *Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court.

Hoyningen-Huene, P. (1992) The Interrelations Between the Philosophy, History and Sociology of Science in Thomas Kuhn's Theory of Scientific Development. *The British Journal for the Philosophy of Science*, 43, 487–501.

Kosso, P. (1992) *Reading the Book of Nature. An Introduction to the Philosophy of Science*. Cambridge: Cambridge University Press.

Kuhn, T. (1957) *The Copernican Revolution. Planetary Astronomy in the Development of Western Thought*, 3rd ed. Cambridge, MA: Harvard University Press, 1970.

Kuhn, T. S. (1970a) *The Structure of Scientific Revolutions*, 2nd ed. Chicago, IL: University of Chicago Press.

Kuhn, T. S. (1970b) Reflections on My Critics. In I. Lakatos and A. Musgrave (eds.) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, pp. 231–278.

Kuhn, T. S. (1977) Objectivity, Value Judgment, and Theory Choice. In *The Essential Tension. Selected Studies in Scientific Tradition and Change*. Chicago, IL: University of Chicago Press, pp. 320–339.

Kuhn, T. S. (1980) The Halt and the Blind: Philosophy and History of Science. *The British Journal for the Philosophy of Science*, 31, 181–191.

Lakatos, I. (1970) Falsification and the Methodology of Scientific Research Programmes. In J. Worrall and G. Currie (eds.) *Imre Lakatos. The Methodology of Scientific Research Programmes (Philosophical Papers I)*. Cambridge: Cambridge University Press, 1978, pp. 8–101.

Lakatos, I. (1971) History of Science and Its Rational Reconstruction. In J. Worrall and G. Currie (eds.) *Imre Lakatos. The Methodology of Scientific Research Programmes (Philosophical Papers I)*. Cambridge: Cambridge University Press, 1978, pp. 102–138.

Lakatos, I. (1973) The Role of Crucial Experiments in Science. *Studies in History and Philosophy of Science*, 4, 309–325.

Lakatos, I. (1978) Anomalies Versus 'Crucial Experiments'. A Rejoinder to Professor Grünbaum. In J. Worrall and G. Currie (eds.) *Imre Lakatos. Mathematics, Science and Epistemology (Philosophical Papers II)*. Cambridge: Cambridge University Press, pp. 211–223.

Laudan, L. (1977) *Progress and its Problems. Toward a Theory of Scientific Growth*. Berkeley, CA: University of California Press.

Laudan, L. (1984) *Science and Values. The Aims of Science and their Role in Scientific Debate.* Berkeley, CA: University of California Press.

Nola, R. and Sankey, H. (2000) A Selective Survey of Theories of Scientific Method. In R. Nola and H. Sankey (eds.) *After Popper, Kuhn and Feyerabend. Recent Issues in Theories of Scientific Method.* Dordrecht, The Netherlands: Kluwer, pp. 1–65.

Quine, Willard V. O. and Joseph S. Ullian (1978) *The Web of Belief.* 2nd ed. New York: Random House.

Salmon, W. C. (1990) Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes. In C. W. Savage (ed.) *Scientific Theories* (*Minnesota Studies in the Philosophy of Science XIV*). Minneapolis, MN: Minnesota University Press, pp. 175–204.

Shapin, S. and Schaffer, S. (1985) *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton, NJ: Princeton University Press.

Worrall, J. (1993) Falsification, Rationality, and the Duhem Problem. Grünbaum versus Bayes. In John Earman et al. (eds.) *Philosophical Problems of the Internal and External Worlds. Essays on the Philosophy of Adolf Grünbaum.* Pittsburgh, PA: University of Pittsburgh Press/Konstanz: Universitätsverlag, pp. 329–370.

Worrall, J. (2000) Pragmatic Factors in Theory Acceptance. In W. H. Newton-Smith (ed.) *A Companion to the Philosophy of Science*. London: Blackwell, pp. 349–357.

METHOD AND OBJECTIVITY

Commentary on "The Aim Structure of Methodological Theory", by Martin Carrier

MICHEL BITBOL

Conceptions of scientific theories are usually distributed into two distinct subsets. The first one is normative and teleological. According to it, scientific theories have or should have "epistemic value"; they have or should have credentials for approaching isomorphism with a putative "external reality" construed as a final target and a criterion of truth; and therefore the ultimate structure of theories is *necessary*. The second one has an evolutionist tinge; it restricts its normative aspect to viability. Here, no epistemic value is required, but only adaptative value; no truth, but empirical adequacy; no pre-defined final target, but a proteiform quest for ecological niches; no necessity of correspondence, but historical contingency. A third conception, a "middle way", can however be identified in the history of ideas. This alternative conception (called transcendental epistemology) was first formulated in a highly fixist version by Kant, and later made more flexible and historically sensitive by the neo-kantian lineage. In this third conception, epistemic value is retained, yet only as a regulative ideal. The claim of truth is no longer discarded but it is thoroughly redefined. Truth is not restricted to logical coherence, nor does it imply mere correspondence with "things-in-themselves". Rather, "objective truth" means "connection according to laws of experience"[1] provided in advance by our understanding; namely connection of phenomena according to those very "constitutive laws" whose application are a condition of possibility of any experience of objects. Moreover, in Cassirer's version of neo-kantianism, the constitutive laws are historically drifting, by way of a progressive conquest of accurate "symbolic forms"; but the content of scientific theories is still ascribed a certain amount of *internal necessity* in so far as it *must* incorporate the group-structure whose invariants define its own objects (Cassirer, 2004).

The *method* of theory choice is then seen in a very different light according to the status one ascribes to scientific theories. Each one of the three former views of theories is naturally associated with a certain approach to methodology. If what is retained from scientific theories is only their adaptative value, methodology is bound to remain extremely flexible and pragmatic. Among the manifold theories that are either merely

[1] I. Kant, *Prolegomena to any Future Metaphysics*, Sect. 49.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 291–296.

empirically adequate, or able to guide a consistent set of efficient practices (according to the underdetermination thesis), the choice can be determined by any set of loose criteria, from mere individual, cultural, social, or political preference, to intersubjectively shared values of a small community of scientists. Even paradigm shift can be decided or delayed along with non-empirical criteria that are dependent on the individual and collective atmosphere of a given moment of the history of science. Kuhn's "big five" non-empirical criteria only happen to be part and parcel of a certain common inheritance of practicing scientists. But, if the accepted final aim of science is to formulate theories that are epistemically credible, or even epistemically *warranted* as *true* or *faithful* pictures of reality, then the role and the status of method become so crucial that no approximation can be allowed. Indeed, it is widely accepted, even by hard-core realists, that the aim of truth or faithfulness can only be reached in the long run. This being granted, one absolutely needs a methodological corpus that inexorably canalizes the scientific investigation towards its final accomplishment, and prevents it from any risk of going astray. Hence the quest for a reliable "methodological theory", such as Bayesianism or Lakatos' progressive research program strategy. The problem is that, as Martin Carrier demonstrates by giving convincing examples in the history of physics or astronomy, no such ultimately satisfactory and "certain" methodological theory appears to be available. The theory of method is exactly as risky and debatable as the scientific theories whose process of discovery it is supposed to direct. The scientific endeavour is open-ended at every layer; not only the layer of theories, but also the layer of the methods that are supposed to lead us towards the edification of theories. And we have no reason to believe that adding a third layer, say a method for formulating methods, would change substantially this situation. Each stage of this regress only renews the dimension of gambling that is typical of science, without giving us any assuredness that the whole process is on the "right" track. We can only take comfort from the fact that this is a very systematic and coherent network of gambling, likely to gain us the advantage of a proper "Dutch book".[2]

After reading Martin Carrier's article, one is then thoroughly convinced that convergent realism is lacking for a proper methodological ground. And, at the same time, one clearly sees that methodological theories are not to be dismissed too quickly, provided they do not claim to offer the undebatable clue for leading science towards its own final closure. Here, we see a methodological "middle way" emerging in the process: acceptance of the value of methodological theories, but denial of the claim that methodological theories are somehow more definitive than the scientific theories they help select. Couldn't one articulate the two "middle ways" that we have just identified, namely Carrier's middle way about method and the neo-kantian middle way about the status of scientific theories? I'll try to show that this is not impossible.

The first point of articulation between the neo-kantian middle way about theories and Carrier's middle way about method, deals with regulative ideals. Although the methodological layer of science cannot claim any absolute certainty, it is just as much attracted by a regulative ideal as the theoretical layer itself. Whereas the regulative

[2] A Dutch book is a sequence of odds and bets which assures a profit, irrespective of the outcome of the game.

ideal of the theoretical layer is perfect mapping of "external reality" yielding maximal efficacy in action, the regulative ideal of the methodological layer is perfect guidance towards this future perfect mapping. These regulative ideals may well be indispensible in order, says Kant, to avoid depriving reason "(…) of its most important prospects, which can alone supply to the will the highest aim for all its endeavor".[3] But it must be clear that they are only *ideals* and by no means representations of the hypothetical aim to be *actually* reached.

The history of methodology displays this ideal status in the most convincing way. Indeed, every important step in theoretical elaboration was first associated with a de facto transformation of method, and later with a retrospective formulation of methodology in normative terms by the very authors of the theories. I'll just give three examples of this process. The example of Newton who deeply redefined the methods of physics in terms of mathematical laws unifying phenomena, and then formulated, in the latest versions of his *Principia* (Newton, 1999), a thoroughly methodological version of his celebrated *Regulae Philosophandi*. The example of Claude Bernard, who made a silent revolution in physiology, and later wrote a treatise of experimental method (Bernard, 1966). And the example of Pierre Duhem (rightly quoted by Martin Carrier), who strongly contributed to establish the independence of the method of macroscopic thermodynamics with respect to the method of mechanics, before he wrote his influential methodological work *La théorie physique*.

The second point of articulation between the neo-kantian middle way about theories and Carrier's middle way about method, is objectivity. According to the neo-kantian school, the major constraint exerted on physical theories is striving towards comprehensive law-likeness and maximal invariance (with respect to individual, spatial and temporal situations), and therefore striving towards maximal objectivity in Kant's sense. Now, if one examines carefully the most widely accepted methodological "ampliative" criteria or values, such as Kuhn's, one finds very similar orientations. Let me just consider three of Kuhn's "big five": consistency, broad scope, and simplicity. Consistency means extensive and tightly packed law-likeness of theoretical prescriptions. As for broad scope and simplicity, they both point towards the ideal of unification under a minimal number of arbitrary assumptions. The ideal of unification, in turn, does for theoretical domains what Kant's constitutive principles do for phenomena: finding invariants. These larger-scale invariants tend towards defining global fields of objectivity, by connecting the local fields of objectivity defined within each particular theory. The universal bearing of group-theoretic structures in intra-theoretical and inter-theoretical architectures provides us with a clear sign that the quest for objectivity is the common ground of many methodological values.

I wish also to point out other signs pointing towards the same conclusion in the history of methodological criteria. The example I will consider bears on Einstein's General Relativity. Why did this theory, with its extensive use of Riemann's geometry, prevail over empirically equivalent alternatives that use other types of geometrical structures? Many reasons have been offered to explain this preference. Some of them just express ignorance of the availability of credible alternatives. But other explanations meet the

[3] I. Kant, *Prolegomena to any Future Metaphysics*, Introduction.

challenge. Let me examine two of them. One is H. Reichenbach's, and the other one is C. Glymour's.

Very early, in a book written in 1928, Reichenbach emphasized the plurality of possible geometries for gravitational theories, and gave a good reason to adopt Einstein's preference for Riemannian geometry (Reichenbach, 1957). As a preliminary move, Reichenbach noticed the crucial role of the background assumptions on measuring *rods* for a proper determination of geometry. One such assumption is that rods must not be *differentially* deformable since, if this were the case, the laws of both physics and geometry would depend on the materials used to make the rods. But no such argument against the restriction of the domain of validity of laws is enough for preventing in principle these rods from being *universally* deformable, under the effect of global forces acting on every bodily element of the world. According to Reichenbach's well-known theorem of relativity of geometry (which was greeted by Carnap as a major result of the philosophy of physics), nothing precludes the possibility that the geometry which arises from systematic length measurements be a purely *apparent* geometry underpinned by a completely different *real* geometry. This may prove true if a universal force is stretching every rod and every measured object as well, in such a way that comparing them can reveal no discrepancy with the *apparent* geometry. For instance, nothing precludes the possibility that the Riemannian geometry used in General Relativity is only an illusory geometry underpinned by a real Euclidean geometry, provided one accepts that there are universal forces whose local intensity is exactly what is needed for explaining *appearances* of curvature in a world that is "really" flat. Several Euclidean theories of gravitation, such as W. Thirring's (1961), have since then instantiated Reichenbach's general claim. Then, according to Reichenbach, physical geometry cannot be construed as something absolute; it only expresses a *relation* between a set of rods (taken as rigid) and the rest of the universe (Reichenbach, 1957, p. 37). At this stage, Reichenbach was aware that he was at risk to accept a form of (geometrical) relativism. To avoid this, he suggested a solution that looks like conventionalism. His solution is called the "principle of elimination of universal forces", i.e., the decision (or convention) to declare that the value of these universal forces is zero everywhere.

Another approach is C. Glymour's. Glymour (1977) fully acknowledges, as Reichenbach did, the plurality of available physical geometries, even in view of the gravitational phenomena specifically accounted for by General Relativity. Yet, according to him, there is no underdetermination of geometry by experimentation *as a whole*. This being granted, there is no need for a conventional principle such as the elimination of universal forces. But how is this possible? How can there be many geometrical frameworks in which gravitational phenomena can be "saved", yet no true underdetermination of physical geometry? Glymour's answer to this question consists in making a crucial difference between "empirical adequacy" and "degree of empirical confirmation". Two theories can be empirically equivalent (insofar as they account for the same phenomena), but have a different degree of empirical confirmation. They have a different degree of empirical confirmation when one of them includes more hypotheses that are not *directly* testable by experiment than the other one. This surplus of non-testable hypotheses, says Glymour, makes the first theory inferior to the second one in spite of its ability to predict the same phenomena.

Now, the non-standard theories of gravitation which, unlike Einstein's, do not use a Riemannian geometry, include a surplus of non directly testable hypotheses. Indeed, they tend to separate the dynamical metric field of General Relativity into: (1) a fixed Minkowskian metric and (2) a dynamical gravitational field tensor depending on the distribution of momentum and energy (Glymour, 1977). Due to the additional degrees of freedom provided by this separation, neither the metric derived from Special Relativity, nor the gravitational field tensor are univocally defined by the Geometrodynamics of General Relativity, from which they have been mathematically inferred. Each one of the geometrical and dynamical elements taken separately appears to be a surplus hypothesis, because only their *combination* is experimentally testable. Therefore the non-standard, Minkowskian, theories of gravitation have a lesser degree of empirical confirmation than General Relativity. Here, the underdetermination of geometry by experiments has been undercut by using a tacit convention that can be characterized as a principle of economy or parsimony: "use a minimum of non directly testable hypothesis".

Let us now think about the status of these conventions (or ampliative criteria) used by Reichenbach and Glymour to suppress underdetermination. Are they truly arbitrary? Or do they both point towards a deeper and common methodological principle? I'll try to show that they can indeed be derived from a deeper methodological principle typical of transcendental epistemology.

Let us start again from the principle of elimination of universal forces. This principle is not so much a statement, as a rule of separation between two disciplines within the architecture of science. What could involve the eliminated universal forces concerns geometry, whereas what involves ineliminable differential forces concerns physics. Geometry incorporates the clauses of reliability and stability of the basic instruments (such as rods) that physics needs in order to confront its theories to experimental phenomena. This does not preclude that certain results of physics exert a decisive constraint in favor of a complete recasting of the axioms of geometry, if this seems to be the only possibility to restore the unity of the physical theories that incorporate geometrical premises. But the distribution of offices between geometry and physics is not altered. During non-revolutionary periods, geometry works as a background presupposition of physics, whereas physics is on the front line, with the task of deciding whether or not the transformation of its constitutive presuppositions is a price to be paid for maintaining its coherence. It then proves quite easy to interpret this dichotomy in transcendental terms: the disciplinary division between geometry and physics is the exact equivalent of the functional division between form and content, or between constitutive principles and constituted objects. This is not to say, of course, that there is no component of content in geometry and no formal component in physics, nor that the formal background of geometry is invariable. But at each step of the evolution of science, the global *role* of form and the global *role* of content must be played, and this is exactly what is allowed by the separation of disciplines tacitly imposed by the principle of elimination of universal forces. This latter principle therefore appears as much more than a mere convention, or one ampliative criterion of theory selection among others: it is a partial realization of one of the most basic preliminary requirements for the act of constituting objectivity.

Similar remarks can be made about the principle of empirical confirmability of the hypotheses of a physical theory. This principle makes explicit a tacit practice of any scientific work: minimizing the number of non-corroborable hypotheses in a theory, or looking for the smaller common nucleus of empirically equivalent theories. But what motivates this tacit practice? It seems clear that the reason why this strategy is almost universally adopted is that it allows to maximize the consensus about theories, thus transforming them into a paradigm. Minimizing the number of non-testable hypotheses is an effective condition for an extended intersubjective agreement about theories, since any such non-testable hypotheses can only be adopted and argued for on the ground of individual or cultural preferences.

To recapitulate, two major ampliative criteria of theory selection (the principle of elimination of universal forces and the principle of maximal corroboration) have a transcendental basis. Both of them express pragmatic constitutive presuppositions of any scientific work. Both of them are underpinned by an even more fundamental criterion of theory choice, which is the quest for ever more extensive domains of objectivity. If science is to define proper objects of study, then it is bound to make use of such principles. This is the component of strong necessity I see in the selection of methodological guides for theory making. This aspect of method can conflict with no other one because it is more basic than any other. It is certainly an excellent candidate for playing the role of a unificatory keystone that Martin Carrier ascribes to methodological theories. And, since it is rooted into one of the most basic definitional features of science, it is likely to remain a permanent aspect of these methodological theories.

So here is, to sum up, the objection I am offering to Martin Carrier's highly cogent analysis of methodological theory. To be sure, the complete process of edification of science, which piles up the formulation of theories and the quasi-simultaneous proposal of methodological principles, is open-ended and without any absolute warrant. I also agree with Martin Carrier that, in spite of this, one should not take methodological theory too lightly because it is part and parcel of the self-imposed teleological and unificatory perspective of the scientific undertaking ("a systematic order and classification"), which is one major reason of its success. However, unlike Martin Carrier, I believe that method is highly constrained. True, it is not constrained from above, namely not by any final formula for ensuring that science will be able to reach its ideal of faithful mapping in the long run. But it is constrained from below, by the very tacit definition of science as a project of objective law-like ordering of phenomena.

BIBLIOGRAPHY

Bernard, C. (1966) *Introduction à l'étude de la médecine expérimentale*. Paris: Flammarion.

Cassirer, E. (2004) *Substance & Function and Einstein's Theory of Relativity*. New York: Dover Phoenix.

Glymour, C. (1977) The Epistemology of Geometry, *Noûs*, XI, 227–251.

Newton, I. (1999) *The Principia (Mathematical Principles of Natural Philosophy)*. Berkeley, CA: University of California Press.

Reichenbach, H. (1957) *The Philosophy of Space and Time*. New York: Dover.

Thirring, W. (1961) An Alternative Approach to the Theory of Gravitation. *Annals of Physics*, XVI, 96–117.

PART 11

A NEW KIND OF INCOMMENSURABILITY AT THE LEVEL OF EXPERIMENTAL PRACTICES?

THE INCOMMENSURABILITY OF EXPERIMENTAL PRACTICES: AN INCOMMENSURABILITY *OF WHAT*? AN INCOMMENSURABILITY *OF A THIRD TYPE?*

LÉNA SOLER

Abstract In the 1990s, authors such as Andrew Pickering and Ian Hacking, who worked on experimental practices and stressed the need for philosophy of science to take into account material and performative aspects of laboratory life, claimed to have discovered a new kind of incommensurability. Pickering talks about a "machinic incommensurability" while Hacking speaks of "literal incommensurability", but their target is approximately the same: an incommensurability related to a disjunction between the sets of instrumental devices, and correlatively the sets of measures obtained by their means, of two scientific practices. The paper discusses this proposal of Pickering and Hacking. It examines the situations that could correspond to the "machinic-literal" incommensurability. It elicits the possible implications of such situations from an epistemological point of view. And it analyses some difficulties which arise from the idea of an experimental incommensurability.

Keywords Incommensurability, Experimental incommensurability, Scientific symbiosis, Contingentism, Inevitabilism, Scientific realism, Weak neutral current.

1. INTRODUCTION

Within post-positivist philosophy of science "incommensurability" is the term used to refer to a particular kind of difference *between theories*, and the "incommensurability problem" consists in the task of characterising the nature and the epistemological aspects of the *theoretical* changes involved. We can, from today's point of view and by contrast, call this characterization of incommensurability "traditional" or "classic". Classically, incommensurability names a relation between elements of the *theoretical* sphere: either a mismatch between theoretical conceptual structures (nowadays commonly called, in the specialized literature, "semantic" or "taxonomic" incommensurability); or an irreducible incompatibility between theoretical standards, norms and values (most of the time named "methodological incommensurability") (see Sect. 4 for more details).

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 299–339.

However, over the past decades, many scholars have voiced their opposition to approaches that focus only on the role of thought and language constitutive of theories, or on the role of conceptual schemes and world views. They have denounced the resulting occultation of an essential dimension, namely that of action, and most of all experimental action. From this enlarged perspective, preoccupied by the importance of performative elements and primarily concerned with the material, instrumental and operational side of laboratory practices, authors such as Andrew Pickering and Ian Hacking have claimed that what becomes apparent is "a new and fundamental type of incommensurability",[1] "that has nothing to do with 'meaning change' and other semantic notions that have been associated with incommensurability".[2] This new type of incommensurability would correspond to the following type of contrast between two experimental practices: the case of two physics based on two completely disjoint sets[3] of measurement instruments and material devices, that is, the case of two physics that would, *literally*, not have a single measure in common.[4] (See Sect. 2 for more details.)

Hacking writes of a "literal incommensurability", Pickering of a "machinic incommensurability". In this paper, I will use the composite expression "machinic-literal incommensurability", so long as it is unproblematic to take Pickering's and Hacking's descriptions as two different names for one and the same reality.[5] I will sometimes also talk about an experimental incommensurability, to capture the more general idea – deployed in Pickering's writings – of an incommensurability that may occur at the experimental level, between ingredients of experimental practices (possibly an incommensurability of a different type than "machinic-literal" ones). If such possibilities were indeed instantiated, the literal-machinic incommensurability would correspond to a special case of experimental incommensurability.

The aim of this paper is to contribute to the debate opened by Pickering and Hacking. The contribution will consist, on the one side in analysing the nature of machinic-literal

1 (Hacking, 1992, p. 54). Quoted and endorsed by Pickering (1995, p. 187). A few lines before he writes that his analysis makes "incommensurability *thinkable in a new and straightforward way*".

2 (Hacking, 1992, pp. 56–57). My emphases.

3 In the most extreme case.

4 A reader familiar with the traditional incommensurability problem might be tempted to protest immediately (as did the referees of this paper) that the relation between two physics deprived of any common measures is definitely not a relation of incommensurability because the necessary element of conflict or competition between the two physics is completely missing. I am well aware of this problem. I provided a preliminary (admittedly still sketchy) discussion of it in (Soler, 2006c), and it is precisely one aim of the present paper to analyse it further. As we shall see, the competition between the two compared scientific practices is indeed a crucial condition with respect to incommensurability verdicts. This condition will play a central role in the discussion, Sect. 6, of the relation between the simple fact of a literal absence of common measure on the one hand, and the judgment that an incommensurability is involved on the other hand. To anticipate, the discussion will lead to the two following conclusions. First, the case of scientific practices that *both* compete and are deprived of any common measures in the literal sense is perfectly conceivable. Second, the condition of an *actual* competition between the two scientific practices under comparison – which corresponds to the condition put forward by those who worked on classical incommensurability – is too strong. A weaker necessary condition is in fact sufficient: the condition of a competition *in the past*.

5 I prefer not to opt for one or the other of Pickering's and Hacking's labels and use one for the other indifferently, since, as we will see, it is not at all obvious that the characterizations in terms of machines and in terms of measurements necessarily refer to one and the same situation. The expression "machinic-literal incommensurability" is admittedly ponderous. But it has the advantage of keeping constantly before our eyes that two different characterizations are originally involved and to remind us that their equivalence does not go without saying.

incommensurability, of its implications from an epistemological point of view, and of some difficulties raised by the idea of an experimental incommensurability; on the other side, in establishing some specific points about experimental incommensurability. More precisely, I will react to Pickering's and Hacking's claims about a new kind of incommensurability along two lines.

First, I will clarify the different situations susceptible to be the basis of a verdict of machinic-literal incommensurability. This is required since, as we will see, the clause "literally no common measure between two scientific practices" can refer to heterogeneous situations *that are not all appropriately described under the heading of "incommensurability"*. "Appropriateness judgments" of the latter kind obviously rely on decisions about what justifies a philosopher of science in describing a given historical configuration in terms of incommensurability. I will develop and defend my own decisions on that point: I will render explicit what I take to be minimal criteria for adequate verdicts of incommensurability. After having explained in what sense these criteria are satisfied for traditional incommensurability, I will turn to Hacking's and Pickering's "new type of incommensurability" and will argue that the clause "literally no common measure between two scientific practices" is, in itself, not sufficient to guarantee incommensurability (according to the stated criteria). I will identify some additional conditions that must be met before being entitled to convert the fact of a literal absence of common measure into a verdict of incommensurability.

Secondly, I will show that one of Pickering's central examples of experimental incommensurability can be re-described as a special case of methodological incommensurability. This alone, of course does not imply that all the historical candidates of experimental incommensurability reduce to cases of methodological incommensurability. But this raises doubts about the novelty of Pickering's experimental incommensurability. Correlatively, this suggests that more attention should be paid to the methodological side of incommensurability than Pickering and Hacking have. Indeed, both of them base their concept of machinic-literal incommensurability, primarily if not exclusively, on a contrast with semantic incommensurability. This also suggests a kind of generalisation, or at least an extension of focus, of the idea of methodological incommensurability: methodological incommensurability could be found, not only at the level of the norms of theoretical appraisal (this corresponds to the traditional focus), but moreover, as well, at the level of the norms of very local experimental practices. Insights will be given about the nature of such a local, experimental methodological incommensurability, about the epistemological issues it raises, and finally, about the relevance of a characterization focused on *local* experimental ingredients.

2. THE IDEA OF A "MACHINIC" OR "LITERAL" INCOMMENSURABILITY

Pickering and Hacking have developed their point of view on machinic-literal incommensurability in close interaction. They regularly refer to one another. Although they describe the "new form of incommensurability" in their own style, they clearly have in mind a similar cluster of situations.

In brief, the idea is the following. If we work out all the consequences of the fact that what plays the role of empirical data, in physics, is constitutively relative to a set of measurement instruments and devices (an *instrumentarium*, as Robert Ackermann calls this set (Ackermann, 1985)), then we are led to the possibility that two physics may be incommensurable in the sense that they would share no instruments, and therefore would have no common measure *in the very proper sense of the expression*.

Hacking calls this new form of incommensurability "literal" (Hacking, 1992, p. 31): as the scientific practices in question have disjoint instrumentaria, they have *literally* no common measure, that is, they have no measurement result in common.

> [D]ifferent and quite literally incommensurable classes of phenomena and instrumentation. I say incommensurable in the straightforward sense that there would be no body of instruments to make common measurements, because the instruments are peculiar to each stable science. (Hacking, 1992, p. 31)

Pickering talks about a "machinic incommensurability" (Pickering, 1995, p. 189), that he specifies as "[T]he disjuncture of the machinic bases of the two regimes, with their differing material performances" (Pickering, 1995, p. 188), or again, as the "shift" of the "machinic termini" (Pickering, 1995, p. 190), the "machinic termini" (or machinic ends) naming the material aspect of what is obtained as the result of the use of measuring instruments.

Pickering and Hacking present their machinic-literal incommensurability as a novelty in philosophy of science and as something radically different from semantic incommensurability.

In Pickering's terms: "We glimpse the possibility of a form of incommensurability that *cannot be articulated within the representational idiom*, an incommensurability in captures and framings of material agency and in the representational chains that terminate in them. This is the *new way* in which the mangle makes incommensurability thinkable". (Pickering speaks, negatively, of a "non representational incommensurability".[6])

Hacking, as already quoted in the introduction, talks about "a new and fundamental type of incommensurability", "that has nothing to do with 'meaning change' and other semantic notions that have been associated with incommensurability". One of his formulations even suggests, about a particular but paradigmatic instance of semantic incommensurability, that the characterization in terms of literal incommensurability should *replace* the characterization in terms of semantic incommensurability: "It used to be said that Newtonian and Relativistic theory were incommensurable because the statements of one could not be expressed in the other – meaning changed. *Instead* I suggest that one is true to one body of measurements given by one class of instruments, while the other is true to another". (Hacking, 1992, p. 54) (my emphases).

Both Hacking and Pickering mention historical scientific situations in which machinic-literal incommensurability is supposed to be instantiated. The central example, articulated in detail in (Pickering, 1984) and subsequently invoked by Hacking, concerns the history of elementary particle physics from the 1960's to the 1980's. For instance, the scientific practices of the "old physics" (the physics of the '60s) and the scientific practices of the "new physics" (the physics of the '80s), taken globally as two wholes, are described by Pickering as incommensurable in the machinic sense.[7] I shall discuss aspects of Pickering's example in detail below.

[6] (Pickering, 1995, p. 188), my emphases; see also p. 190 and 192; p. 223, n. 14.

[7] The expression "machinic incommensurability" does not appear in (Pickering, 1984) in which the example is first deployed in detail. But in (Pickering, 1995, p. 189), the example is re-described in this vocabulary: "The transformation from the old to the new physics is an example of the kind of machinic incommensurability that the mangle leads us to expect".

3. MINIMAL CONDITIONS FOR A VERDICT OF INCOMMENSURABILITY

The task of discussing Pickering's and Hacking's literal-machinic incommensurability has two faces. On the one hand, we must clarify what kinds of scientific configurations may be concerned (that is, which situations may correspond to the clause "no measure in common" between two scientific practices). On the other hand, we must discuss if and why each of these possible configurations is, or is not, adequately described *as a case of incommensurability*. With respect to this second aspect, we need to explain our decisions regarding the conditions under which a description of a situation *in terms of incommensurability* is "appropriate", "justified" or "legitimate". In this section, I will present, and justify as far as possible, my own decisions on the matter.

3.1. Two minimal conditions that should be fulfilled in order to legitimate the description of a given situation in terms of incommensurability

When one labels two entities (theories, practices) "incommensurable", one means, at least and in first approximation, that these two entities show essential differences that make them difficult to relate to each other. In my opinion, two minimal requirements should be satisfied in addition for a legitimate use of the term "incommensurability". I will first state these two conditions. Then I will explain the second one in more detail and will provide elements of justification for it.

(a) First requirement: to be able to do justice to the seminal idea of a lack of common measure; in other words, to be able to specify the kind of common measure that is missing. I will call this requirement the ACM (Absence of Common Measure) condition.
(b) Second requirement: to be able to do justice to the "catastrophic" connotations of the term "incommensurability". I will call this requirement the C condition (C for Catastrophic).

By "catastrophic connotations", I refer to what I take to be a fact – a sociological fact about the group of people who reflect on the nature of science: for most of these people[8] the term "incommensurable" is, almost inevitably and automatically, associated with the pretension that something essential is questioned about science; as a consequence, the term induces strong reactions; it has a strong emotional potential.

[8] The C condition of course does not hold "in the absolute" but *for a given subject*, specified here as "the group of people who reflect on the nature of science". This group comprises philosophers of science, sociologists of science, historians of science, scientists interested by reflexive studies about science, and so on. Admittedly, the boundary around such a group is not sharp. Admittedly, any claim about the position, attitude, feeling … of this group *considered globally* will have the status of an 'average claim'. In other words, it will almost always be possible to find individual exceptions. But this does not disqualify the soundness of such global characterizations. Actually we cannot escape such blur – unless we renounce saying anything about the intellectual position of a human group as a group.

"Doing justice" to the catastrophic connotations of the term "incommensurability" requires giving a definite philosophical content to this vague and immediate feeling induced by the word "incommensurable". Put differently, one should not label a given epistemological configuration "incommensurable", unless one is able to elicit what, precisely, is questioned by the existence of this configuration, that is, what, precisely, are the damaging consequences for (at least one) well-identified conception of science.

Before giving elements of justification in favour of the C condition, let me analyse point by point what it means.

3.2. Explaining condition C

1. The C condition starts from the *connotation* of the word "incommensurability", and not from some specific well-articulated conception of incommensurability (such as Kuhn's or Feyerabend's conception). It starts from the effect of the word. (There are, of course, some relations between these conceptions and these effects, but this is not the relevant point here.) Hence the "automatically": it is intended to suggest the immediate, unreflective character of the association involved: something like a Pavlovian reflex. This is typical of the way the connotations of any word act on the members of a linguistic community. Now, what is the effect of the word "incommensurability" on the group of people who reflect on the nature of science (let us say, for short, the "reference group")? Beyond individual fluctuations we can say, globally, that the term very largely induces strong, if not violent reactions. These reactions can be negative or positive. But whatever their quality may be, the term "incommensurability" is not neutral. Be the feeling positive or negative, it is not indifference. For that reason, the decision to use the term "incommensurability" rather than another one is itself not neutral. At least, this decision should be carefully considered. This is the main motivation for the C condition.
2. The C condition likewise specifies to some extent the origin of the strong reaction commonly induced by the word "incommensurability": "the pretension that something essential is questioned about science". This clause defines the "catastrophic connotations". What, precisely, is questioned? This does not matter as long as we are concerned with the level of connotations. There could be important variations in the answer that different members of the reference group would give should one ask them to reflect on the matter. At the level of connotations, however, the important thing is that most of these people associate the word incommensurability to the idea *that* something essential is questioned about science. At this level, similarly, the adjective "essential" only refers to the fact that most people of the reference group share the following feeling: what is questioned (whatever it may be) by the incommensurability thesis, is *taken as* a fundamental feature of science *by a non-negligible sub-group of the reference group* (in other words there is, for an important sub-group, a strong association between the idea of science and the questioned feature). For the sub-group who believes that the "questioned feature"

is indeed characteristic of science, the incommensurability thesis must be rejected as erroneous and possibly dangerous, and the connotations of the term "incommensurability" are negative. For the sub-group who do question the association between this feature and science, the incommensurability thesis is right and possibly liberating, and the connotations of the term "incommensurability" are positive. Nonetheless, all share the feeling that something essential is at stake. And this is the important point with respect to our C condition.

It is indeed to that shared feeling that the adjective "catastrophic" refers in the "catastrophic connotations" mentioned in our C condition. The "catastrophic connotations" are not necessarily negative: they may be positive as well. In this expression, the term "catastrophe" should be understood by analogy with the sense involved in mathematical Catastrophe Theory: the sense of a rupture within a given space. In a given space of philosophical positions about science, a rupture arises, in the sense that something important, largely taken for granted, is questioned. The reaction associated with this rupture is, in details, potentially variable from a member of the reference group to another, since it is a function of the relation, in itself heterogeneous, between each single subject and a certain ideal of scientificity. But for all parties the connotations of the term "incommensurability" are "catastrophic", in the sense that all of them consider what is at issue as crucial, that is, as bearing important implications.

3. Finally, the demand to do justice to the catastrophic connotations of the term "incommensurability" corresponds, for the philosopher of science, to the following task: to be able, for each incommensurability verdict, to explain what lies behind the immediate and vague feelings associated with the level of connotations; that is, to succeed in showing that the matter of fact involved has indeed crucial implications for a specified group of people; or in other words, to identify at least one conception of science that is questioned and to elicit what are the related damaging consequences for it.

Having explained the content of condition C, I will now explain why I think it is required.

3.3. Justifying condition C

One could be worried by the fact that according to the decisions above, the connotations of the word "incommensurability", as well as the damaging consequences of the ACM involved, are taken as criteria of incommensurability verdicts *while they are not intrinsic features of the scientific configurations under study*.[9] I agree that obviously, they are not. Indeed, the connotations of a word depend on the contingent way this word has been used in particular contexts of discussion, and this may vary whereas the matter of fact named by the word remains the same. Similarly, the "damaging consequences" of a given scientific configuration are not damaging just because of

[9] Worries of this kind have indeed been expressed by readers of preliminary versions of this paper. This section aims at answering to them.

some intrinsic features of this configuration: they are damaging *with respect to some existing contingent human beliefs about science* they are claimed to question. I also agree on the importance of careful distinction between what pertains to the level of the intrinsic or internal characteristics of the scientific configurations under scrutiny, and what pertains to the level of external characteristics of the human "reference groups" that we take into account. Actually, it is precisely because of the importance of such distinction that I have imposed *two separate* conditions – the ACM condition and the C condition – on judgments of incommensurability.

The ACM condition addresses the level of the intrinsic characteristics of the scientific configurations involved. Indeed, it specifies the kind of ACM that is involved (a semantic one? A methodological one? A machinic-literal one? Another kind?). It describes the matter of fact under study, it defines the type of scientific configuration under scrutiny. If this matter of fact is named an incommensurability, the ACM condition will correspond to the definition of this incommensurability.

Now, in the presence of a well-characterised type of ACM, we still have to decide in which terms – with what kind of vocabulary – we are going to describe the situation. We still have to decide whether or not it is appropriate to characterize the situation in terms of "incommensurability". It is at this level that the C condition enters the scene. The C condition expresses criteria for this terminological adequacy. And such criteria are this time related, not only to intrinsic features of the scientific configurations involved, but also and primarily, to certain linguistic uses and historical commitments about science, that is, to elements that pertain to the external level of the "reference group".

Having clarified this, what can be said to justify the choice of the criteria at this level? Obviously, in this case as in any discussion about the appropriateness of a terminology, convention plays a role – with the consequence that we can justify our decisions only up to a point. But this obviously does not imply that such discussion, and the terminological decisions finally adopted, are nothing but conventions, are pure questions about words.

Let me begin with connotations. Why should we take connotations into account?

Because most of the time, words are not neutral. They are not just indifferent sounds. This is, of course, because they are historically loaded. They are such, first because they have an etymology (which may be more or less "visible", more or less "audible" in the present, for a given community of language, depending on the subsequent history of the word after its first introduction). And they are historically loaded, second because their effective use has a determinate history that (most of the time) can be traced and leaves a deep impression behind. Conflicts may arise between the etymology and the posterior usages: the actual use is sometimes unfaithful to the etymology. But in any case, the history of the use shapes the connotations that a word carries with it[10] (all these judgments of course do not hold "in the absolute" but *for* a human group that the analysis has to specify).

Now, a philosopher of science who faces the problem of terminological choices should not ignore these aspects. This is especially important for words that carry strong (positive or negative) connotations, that is, for words that have a strong emotional power, as is the

[10] In Sect. 3, I will briefly consider the history of the term "incommensurability" from its first introduction into the field of philosophy of science, in order to understand the origin of its catastrophic connotations.

case for the word "incommensurability". These are words that, when employed, quasi-inevitably and immediately induce violent reactions. If you describe physicists' work as an "irrational" enterprise ultimately driven by "social factors", you should not ignore the fact that most realist philosophers and most scientists will immediately react negatively. They will at least reject your characterization, if not consider you as an enemy of science. (Recall the "science wars". Of course, a symmetric claim could be stated from the other side.) This may be a strategy, for example if your aim is to become famous in Science Studies. (Many writings will mention your name as the enemy to fight.) But if the aim is to understand science, to give a sound characterization of the kind of enterprise it is, and to create the conditions of a real debate with differently oriented colleagues, this is certainly not the appropriate terminological choice. Maybe what you have in mind under "irrational" and "social factors" could be accepted, under scrutiny, by realist philosophers and scientists. But not under this terminology. At least such a vocabulary will create a barrier from the beginning. Few will consider your work in a charitable perspective, few will be in a good disposition to study the details of your analyses.

All in all, I think we should acknowledge the non-neutrality of words and take their effects into account. This motivates and justifies one aspect of my C condition. The second aspect is related to the particular content of the connotations involved in the particular case of the word "incommensurable", namely to the idea that something essential is questioned about science, and it leads to the demand to elicit the damaging implications, for at least one conception of science, of the kind of ACM under scrutiny.

Why should we take these damaging implications into account?

Because if the kind of ACM involved was completely harmless (if it did not indeed question essential aspects of at least one conception of science), and if one nevertheless decided to describe this ACM as an incommensurability, this would create a tension between the catastrophic connotations of the word on the one side, and the innocuous aspect of the ACM involved on the other side.

If, for example, Hacking's literal ACM only named the situation of two distinct disciplines as physics and sociology, everybody would agree that this kind of ACM is exemplified in the history of science but that this scientific configuration has no damaging implications.[11] In such a case, it would not be judicious to speak of an incommensurability – *although one could perfectly well argue that it is appropriate to talk of an absence of common measure in the literal sense*.

Let us now examine *in what sense* the traditional incommensurability satisfies our two requirements, before examining whether the so-called machinic-literal incommensurability satisfies them.

4. THE SENSE IN WHICH TRADITIONAL INCOMMENSURABILITY SATISFIES THE ACM AND C REQUIREMENTS

Since the introduction of the idea of incommensurability by Kuhn and Feyerabend in 1962, post-positivist philosophy of science has gradually identified two kinds of incommensurability (Hoyningen and Sankey, 2001; Soler, 2000, Chap. VII; Soler, 2004).

[11] I will discuss this kind of situation in Sect. 6.2.1.

These two kinds are most of the time intertwined in practice, but they are different in character and must be carefully distinguished from an analytical point of view.

- The incommensurability of theoretical contents, commonly called *semantic* incommensurability.
- The incommensurability of the norms governing theoretical elaboration and appraisal, commonly called *methodological* incommensurability.[12]

To the question "what is incommensurability?", the classic answer is, in its most general formulation: an incompatibility irreducible to a logical contradiction, which arises either at the level of scientific language or at the level of scientific standards, and which appears astonishing and potentially significant from the epistemological standpoint, for it happens to arise between two rival theories and theoretical practices for which everybody would expect a common measure of such a kind.

Granting this definition, classical incommensurability satisfies the ACM condition in the following sense:

- As to semantic incommensurability: there is a lack of common measure at the level of semantic resources (at the level of what can be expressed and symbolized). Taxonomies cannot be superimposed, concepts are not translatable.
- As to methodological incommensurability: there is a lack of common measure at the level of the norms of scientificity that govern theory appraisal.[13] There are no (or too few) shared standards and values.

What about the satisfaction of the C condition? Let us begin to consider the situation from a historical perspective. This will allow us to understand how the catastrophic connotations have been attached to the word "incommensurability".

When Kuhn and Feyerabend first introduced the incommensurability thesis in the 1960s, the pretension that something essential had to be questioned in common conceptions of science explicitly accompanied their characterization. The conception of science they primarily attacked was the "received view", that is, the one of logical positivism (or their reading of it). In the next period, the incommensurability thesis induced a violent controversy about the nature of science. At the most general level, the two main issues were, according to all, realism and relativism.[14]

True, after the first introduction of the incommensurability thesis in 1962, a number of philosophers of science argued that the incommensurability thesis was, after closer examination, not so harmful philosophically, if not completely innocuous, notably with respect to realist and anti-relativist conceptions. However, not all the people reflecting on the nature of science agreed, and even today, this is still not a consensual point.[15] The catastrophic connotations continue to be glued to the

[12] Giere also employs the expression "value incommensurability". (See his commentary on Teller's paper in the present volume).

[13] In Hoyningen's and Sankey's formulation: "there are no shared, objective methodological standards of scientific theory appraisal. Standards of theory appraisal vary from one theory or paradigm to another" (Hoyningen and Sankey, 2001, p. xiii).

[14] See the references given at the beginning of this section. (See also Sankey, 1994).

[15] For example, philosophers of science such as Howard Sankey or Alexander Bird argue that incommensurability is compatible with realism (see for instance, Sankey, 1998; Bird, 2000), whereas Paul Hoyningen-Huene argues the opposite. (See his commentary on Bird's paper in this volume).

term "incommensurability". The term continues to work, beyond and sometimes largely independently of the matters of fact Kuhn and others after him wanted to pick out with this label, as a kind of abbreviation for the claim that something essential to our conception of science is being questioned.

As a philosopher of science whose research program has been explicitly centred on the "incommensurability problem" for many years, I had the occasion to experience that still nowadays the term "incommensurability", even followed by the word "problem", induced many violent (in my experience almost always negative) immediate reactions, independently of any examination of the work articulated under this heading.

To switch from the historical perspective to an analytical characterization and to be more precise, classical incommensurability satisfies the C condition in the following sense:

- Semantic incommensurability casts doubts on the descriptive value of science, and this is potentially damaging for realists.
- Methodological incommensurability seems to imply relativism, as the thesis that scientists' decisions are ultimately determined by factors that are subjective, social, or, at any rate, external to the reality under study. This is damaging for foundationalists, realists, inevitabilists[16] and for all those who see scientific rationality as a universally compelling power.

Now, what about machinic-literal incommensurability? In what sense does it satisfy the ACM condition?

5. WHAT IS (OR COULD BE) A MACHINIC-LITERAL ACM?

The situations that Pickering and Hacking describe as cases of literal-machinic incommensurability satisfy the ACM condition in the sense that the two scientific practices under comparison do not have any instrumental device in common and do not have any physical measure in common (Absence of Common Machines and Absence of Common Measures). At first sight, this seems a rather clear and unproblematic clause. But is it really? The present section will suggest that it is not and will identify some

[16] The term "inevitabilist" is introduced by Hacking (see Hacking, 1999, 2000). Inevitabilists are those who believe that, under the empirical condition of a progressive physics, physics had inevitably to resemble to ours, at least as far as the main results are concerned. The opposing position corresponds to "contingentism". According to contingentists (like Pickering), science is irreducibly path-dependent: the results are not separable from the process. At each stage, the stable results that possibly appear are the emergent products of the contingent factors that were in place at an earlier stage. So that we could have ended up with a science at the same time *as efficient and progressive as ours*, but *with no common measure*, *at any level*, with ours: for example a science making no use of quarks or of Maxwell's equations, and in which completely different elements would play the role of basic established "empirical data". For further discussion, see (Soler and Sankey, 2007). This is a symposium that includes an introduction by Léna Soler, "The Contingentism versus inevitabilism issue", and four contributions: L. Soler, "Revealing the Analytical Structure and some intrinsic major difficulties of the contingentist/inevitabilist issue". Allan Franklin, "Is Failure an Option? Contingency and Refutation". Emiliano Trizio, "How many Sciences for one World? Contingency and the Success of Science". Howard Sankey, "Scientific Realism and the Inevitability of Science". See also (Soler, 2006a) and (Soler, 2006b).

of the difficulties involved. In order to bypass them and to make progress in specific issues, I will then introduce a distinction between two regimes of investigation for the philosopher of science: the "opaque" and the "transparent" regimes.

5.1. An analysis of the machinic-literal ACM

Following Pickering's and Hacking's nominal definition of their machinic-literal incommensurability, the ACM here involved corresponds to a disjunction between two instrumentaria and two sets of related measurement results. However, as soon as one tries to give a precise content to this nominal definition, one faces some difficulties.

1. Let us first follow Pickering's idiom: let us start from the "machines" involved in experimental practices and reflect on the idea of a "machinic disjunction".

Every experimental discipline that has attained a certain degree of sophistication involves the use of a variety of instruments (scales, microscopes, etc.). What does lead us to regard a variety of instrumental devices and "machinic performances" as a *unitary* instrumentarium or, instead, as several disjoint ensembles?

The question is far from being trivial. It raises the issue of the identity of an instrumental device. Now, if we ask "what is a measurement instrument?", we can answer along several lines:

– A *material* physical object that yields reproducible traces or material ends in similar physical circumstances
– A *functional* object (an object *for*… measuring, determining, testing… Where what follows the "for" almost always implies some elements of the high-level theories)
– A *conceptual* object referred to a theory of the instrument
– A *handled* object implying know-how and some specific skills

This short list, which is not complete and whose different elements bear complex relations to one another, enables us to grasp the nature of the problem. The measurement instrument is a complex object that lies at the intersection of a variety of dimensions involved in scientific practices: starting from the most basic level of the tangible object, considered in its material aspect, we are, little by little, led to take into account an open-ended and increasing series of new characters, related to different dimensions that all play a role in its fundamental mode of being: concrete actions, know-how, intentional acts, projects, more or less explicit local theories of all kinds.… And moreover, up to a level to be determined case by case, some high-level theories.

Because of the complex and heterogeneous nature of the instrumental device, the question of the identity/difference between several instruments or sets of instruments has no immediate and obvious answer. It calls for a reflection devoted to what principles lie behind the constitution of instrumental classes of equivalence, for an analysis of the criteria that govern the individuation of instrumental kinds.

2. Similar difficulties arise as well if we follow, rather than Pickering's idiom in terms of "machines", Hacking's idiom in terms of experimental "measures" associated with the possibility of a literal ACM.

What, exactly, is a measure? It presupposes an instrumental device, and hence all the dimensions elicited just above. Moreover, it adds the idea of a *result* obtained with the instrument involved. Now, how should we understand the nature of this "result"? What, exactly, is obtained with an instrumental device, that we should identify with the "measures" involved in Hacking's clause "literally, no common measures"?

– Observations of the most basic kind associated with "observational statements" that almost everyone, including the layman introduced into the laboratory, would be able to utter? (For example deviations of scales, plots on screens, photographic plates obtained with bubble chambers.… I will refer to this level of data, either as "machinic ends", using Pickering's terminology (see the beginning of Sect. 2), or as experimental "marks" using Hacking's).
– Data extracted from preliminary analyses? (e.g., data resulting from statistical treatments; or events remaining after the exclusion of ambiguous marks on photographic plates issued from bubble chambers; and the like).
– Attributions of properties (numerical value or qualitative features) to some physical variables? (For example attribution of a quantitative rate for the weak neutral current compared with the charged one).

The answers to the preceding questions are not at all obvious. A fortiori it is not obvious what the answer is to the higher-level question: what is it exactly, that leads us to claim that several measures are *the same* or *different* measures? To claim that two sets of measures pertain to one single unitary set or, rather, to two disjoint sets? Admittedly, the answer may vary according to the meaning of "measure" that we select in the list just above.

All in all, the problem is ultimately the same, whether it is framed in terms of measures or in terms of experimental machines: it concerns the principles on which rest the constitution of equivalence classes (of measures or of instruments); in other words, it concerns the criteria that govern the individuation of identical/different kinds (of measures or of instruments).

3. Moreover and finally, it is not at all obvious that Pickering's formulation in terms of machines on the one hand, and Hacking's formulation in terms of measures on the other hand, are necessarily equivalent in the sense that they necessarily capture *one and the same* class of situations. Do not practitioners commonly say that the same measures are obtainable with the help of different instruments (e.g., that measures of temperature can be achieved with different thermometers)? It seems, thus, that we can give a sense to the idea of "different instruments but same measures". This being admitted, an instrumental disjunction will not necessarily imply a disjunction at the level of measures, and hence, a machinic ACM will not necessarily imply a literal ACM.

5.2. "Opaque" versus "transparent" regimes of investigation

The difficulties and uncertainties put forward in the last section call for a deeper investigation that should lead to proposed solutions. But this would require another paper. Fortunately, even deprived of such solutions, we are not blocked: there remains a way to make progress without giving precise answers to the previous questions. We can, up to a point, bypass the difficulties just mentioned.

We can avoid answering questions such as "which kind of differences are behind the judgments 'identical/different kinds of machines or measures' ", by retreating to the level of practitioners' positions taken as an unquestioned fact (here practitioners' positions regarding whether, in this or that particular historical situation, these two sets of instruments or measures pertain to one and the same unitary set or are two different sets deprived of intersection). In other words, we take practitioners' positions as a given fact and as primary data, without asking what commitments or reasons are behind such judgments for those who endorse them.

I will use the expression "opaque regime" to refer to this way of approaching the problems. In the opaque regime, the judgments (here the judgments about instrumental disjunctions and the like) refer to a specified (collective or individual) subject. Providing that there is a consensus of sufficiently high quality concerning the matter involved, this subject will be a reference group like "most physicists", "most specialists of this domain", and the like. In the opaque regime the judgments of the reference group have the status of sociological facts. These facts are taken into account *as such*, and their possible consequences for scientific practices are analysed *without asking why* the reference group endorses them or if the reference group is right to do so. In brief, in the opaque regime we rely on an un-analysed sense of some actual, practice-based positions. We treat these judgments as data and we discuss their possible implications.

By contrast, a "transparent regime" would correspond, at least to an analysis of the assumptions or criteria that are behind the judgments involved, and possibly, moreover, to a discussion of the reasons why such judgments have been endorsed historically. A "transparent approach" of the experimental disjunction would require a detailed analysis of what constitutes the judgements of experimental disjunction in real historical case-studies or in principle. A transparent approach would attempt to propose answers to all of the questions raised in Sect. 5.1. This would certainly lead to a distinction between different kinds of experimental ACM, according to the different experimental features to which the experimental disjunction is referred. The analysis proposed in Sect. 5.1 suggests, for example, the following possibilities:

– An experimental ACM related to the theories of the instruments: here the machines (and the marks obtained by their means) would be divided into two disjoint sets because their conception would involve different, independent theories, if not contradictory or semantically incommensurable theories. In that case, the *kind of difference* at stake would not be a new kind: it would be a difference between theories, that is, a kind of difference already considered and well-characterized by traditional approaches devoted to the comparison of different stages of scientific development (even if the theories involved here do not necessarily equate with high-level "universal" theories). Switching from the level of the ACM judgments to the level of the incommensurability judgments, we would have to discuss whether or not the experimental situations that might instantiate such theoretical differences at the experimental level can have catastrophic consequences.

- An experimental ACM related to material and geometrical differences: here the two sets of machines could fundamentally involve the same theoretical principles (They could be the *same type* of machines), but a new material would be used for their construction, or an important change of dimension would be introduced, and *because of reasons of that kind*, the two sets of machines, and the measures obtained by their means, would be seen as two disjoint sets. As a possible illustration, we can invoke the bubble chamber neutrino experiments of the 1960s and of the 1970s as Pickering describes them: although both sets of experiments involve the same types of ingredients (neutrino beam, visual detector and the like), the Gargamelle bubble chamber used in the 1970s is much bigger than all its ancestors, and this could be a reason to conceive the 1970s experiments as disjoint from the 1960s experiments. (This kind of situation seems to correspond to what Pickering has very often in mind in his developments of the idea of experimental incommensurability.) Once again, switching from the level of ACM judgments to the level of incommensurability judgments, we would have to discuss whether or not the experimental situations that might instantiate such geometrical-material differences at the experimental level can have catastrophic consequences.
- An experimental ACM referred to differences in the skills and know-how: here two sets of measures would be divided into two disjoint sets because they would be obtained by people who do not master the same skills and know-how. As a possible illustration, we could invoke the well-known episode of Galileo and his telescope: Galileo's measures and, say, Kepler's student's measures with the same telescope at the same place would correspond to two disjoint sets because the two men are differently trained.[17]
- An experimental ACM referred to experimental standards: here a collection of experimental ingredients would be divided into two disjoints sets because they would be differently valued by two scientific communities. I will discuss an example below, in Sect. 7.1.

A transparent approach would have to discuss the autonomy of such different possibilities and the way in which they are concretely intertwined in the constitution of judgments about the unity/multiplicity of instrumental sets. It is only once that work would be accomplished that we could examine the novelty of Pickering's and Hacking's proposals. But even deprived of a precise answer to these questions, we can nevertheless progress in the general problem of an experimental incommensurability, thanks to the distinction between opaque and transparent regimes, for that distinction entitles us to circumscribe modules of sub-problems that can be treated independently.

[17] See the analysis of (Feyerabend, 1975), Chap. 10. We can note that in such a situation it does not seem very appropriate to talk of a disjunction between *machines* – or it would be an unusual, non-intuitive sense of the term "machine".

6. FROM THE FACT OF AN EXPERIMENTAL ACM TO THE VERDICT OF AN EXPERIMENTAL INCOMMENSURABILITY

Let us suppose that we indeed have (in the opaque regime) a machinic-literal ACM between two experimental practices EP1 and EP2. Suppose, moreover, that the definition "no intersection between the sets of machines and the sets of measures" of these practices points to a type of ACM indeed irreducible to the two traditional semantic and methodological types. This would of course require a confirmation in the transparent regime. But suppose it is indeed the case: suppose we have identified a new kind of ACM. Should we conclude, with Pickering and Hacking, that we have discovered a new kind of *incommensurability*? Not yet (see Sect. 3). It remains to examine whether (and if yes in what sense) the C condition is satisfied.

In this section, I will argue that the very *fact* of an experimental *ACM* is not sufficient to generate a genuine *incommensurability*. I will identify two additional necessary conditions, a negative one (Sect. 6.1) and a positive one (Sect. 6.2). Then I will discuss some problematic features that should be possessed by experimental practices in order to transform an experimental ACM into an incommensurability (Sect. 6.3).

6.1. A negative condition that should be satisfied in order to convert an experimental ACM into an experimental incommensurability

Suppose that in a given historical situation, physicists are indeed led to describe, intuitively and massively, two sets of experimental items as two disjoint sets. Since we are in the opaque regime, we take this as a fact, without asking *why* it is so and *what specific features of the two compared experimental practices are responsible* for such a way of thinking. Now, whatever may be the specific differences between the two experimental practices under comparison, and *however strong* these differences may be, *other factors*, and factors that pertain *to the theoretical sphere* and not to the experimental one, influence in a *decisive way* practitioners' judgments about experimental disjunctions. This is the point I want to articulate now.

Consider two rival experimental practices, say EP1 and EP2, that are associated with two disjoint instrumentaria and have, *de facto*, literally no measure in common (once again: these are judgments in the opaque regime). For instance, consider two experimental traditions in high energy physics, the "visual" one and the "electronic" one, both concerned with the question of the constitution of matter at the microscopic level.[18] Or imagine a parallel physics on a twin-earth, involving a set of instrumental devices entirely unknown to our own physics. At a given stage of the history of science, the situation of EP1 and EP2 with respect to high-level theories may be diverse: for example, EP1 and EP2 may be explained by two incompatible (contradictory or semantically incommensurable) theories (*either* T1 *or* T2); or by two different but not irreconcilable theories (T1 *and* T2); or else, there may be a unifying theory T that accounts for the machinic ends of EP1 and EP2.

[18] See the analyses of Galison (1997).

These different situations with respect to high-level theories load the practitioners' judgment of an experimental disjunction with significantly different senses. Indeed, it is only in the case where T1 and T2 appear irreconcilable (*either* T1 *or* T2), that the experimental disjunction can be understood in the strong sense of a *mutual exclusion* between the means and results of the two experimental practices. If, to the contrary, a single unifying theory T is available, or if the two different theories T1 and T2 prove reconcilable, the "disjunction" of the two sets of experimental devices and measures will be a "disjunction" only in a weaker sense. What exactly is this sense? Admittedly, it may vary a little bit in details, according to the various kinds of situations that can exemplify the case of two experimental practices explainable by "a single unifying theory T" or by "two different but reconcilable theories T1 and T2". But beyond these possible variations, this sense is weak since it corresponds, at a certain level at least, to *compatible* differences.

For example, if we consider the fictitious twin-earth case, the experimental "disjunction" will only mean that the experimental devices and results involved on twin-earth were *simply not known* on earth before the reunion of the two parallel scientific trajectories. In the transparent regime, what this "not known" means exactly should be spelled out. But with respect to our present aim, it is enough to stress that whatever the unanalysed "not known" may mean in detail, it clearly goes with the idea that the two primarily separated and at first sight very different (disjoint) experimental practices finally proved *compatible, additive* with respect to their results. Hence the "disjunction" only means a spatio-temporal contingent separation and not a fundamental antagonism.

If we turn to the real case of the visual and the electronic experimental traditions in high energy physics after 1974, we find something structurally similar. After 1974, physicists considered that the existence of the weak neutral current had been experimentally established by two different *independent* experiments: on the one side by experiments using visual detectors, on the other side by experiments using electronic detectors. This is a very common way of talking, one that points to a possible example of an experimental disjunction in the weak sense: here the experimental disjunction corresponds to the "independence" of the two kinds of experiments. In the transparent regime, one would have to analyse the "independence" that is vindicated in such cases. But whatever this independence exactly means, whatever the underlying experimental differences may be, these differences can be named "disjunctions" only in a weak sense as soon as the results of the two experimental traditions prove to be explainable by the same theoretical framework or by "additive" frameworks.

We can also consider the same point in a diachronic perspective. Suppose that at a given moment *t* of our history of science, all physicists actually conceive the instrumentaria and the machinic ends of two experimental practices EP1 and EP2 as disjoint *in the strong sense of a essential separation*.[19] Suppose moreover that subsequently, at a time t', a *unifying* high-level theory comes into play, a theory able to account for

[19] Once again, this is a characterization in opaque regime. In transparent regime the justification will most of the time be related to the fact that practitioners consider that the two experimental practices are responsible for two different kinds of physical realities.

the machinic ends of EP1 *and* EP2 *altogether*. How will the actors react to such a situation? These actors, who at the beginning conceived the machines and the instrumentaria of the two experimental practices as disjoint, will subsequently conclude that the initial apparent experimental disjunction simply amounted to a factual, provisional and anecdotal separation *and not to a genuine mutual exclusion*. They will trace back the initial judgment of a disjunction to some *contingent and extrinsic* reasons linked with the human vicissitudes of the investigation, rather than to compelling factors having to do with the proper content of experimental practices – factors most of the time ultimately related to intrinsic characters of the object under investigation. As in any theoretical unification, they will conclude that what has first wrongly been taken as two distinct specialties in charge of two separate realms of phenomena actually corresponds to one and the same specialty responsible for a single homogeneous realm of phenomena.

Conclusion: Whatever the intrinsic differences between two experimental practices may be, however strong they may appear considered in themselves independently of high-level theories, the classes of machinic ends resulting from these two practices will be considered, at a certain level of interpretation, as pertaining to a unitary instrumentarium, given that they turn out to be explainable on the basis of a single high-level theory. In each stage of the history of science, the identity (or at least the addability) at the theoretical level, when it holds, supplants all the differences that may exist at the experimental level considered in itself. In this specific and narrow sense, there remains a residual pre-eminence of high-level theories over experimental practices.

These reflections encourage us to distinguish two senses of "experimental disjunction". The strong sense corresponds to a mutual exclusion: the experimental differences at issue are conceived as irreconcilable, as essentially incompatible. The weak sense corresponds to differences that are recognized as significant differences between experimental practices but that nevertheless are, at a certain level, reconcilable (reconcilable in fact or thought to be reconcilable in principle even if they are not yet actually).

In situations where two experimental practices prove explainable by one unitary theory, it is the weak sense that is involved. One could protest that to talk of a "disjunction" in such cases is spurious. One could still more convincingly argue, switching from the characterization in terms of disjunction to the characterization in terms of ACM, that it is completely inappropriate to describe the relation between two experimental practices explained by one unitary theory *as* a literal *absence of common measure*, since there is a sense of "measure" in which the measures of the two experimental practices are "common". Indeed, even if the experimental means of two traditions, and the raw experimental marks obtained with them, are very different in important respects, all these raw experimental marks obtained with these experimental means are seen as the sign of the same physical phenomena at a higher level of interpretation, and we commonly name "measures" what is at stake at this intermediate level of interpretation. The problem, here, is induced by the ambiguity of the term "measure" (see Sect. 5.1) and by the fact that Hacking did not specify the sense he referred to when he talked of a literal ACM.

Now beyond such ambiguities and whatever our terminological decisions may be about the right application of expressions like "experimental disjunction" and "literal ACM", one thing is certain: even if we are ready to talk of a literal *ACM* in such situations, we should not, switching from ACM verdicts to incommensurability verdicts, infer a literal *incommensurability* between the experimental practices at issue. Indeed, if the experimental differences can be passed beyond at the level of what is taken as experimental results, what could be the potential damaging consequences? I cannot find any.[20] Hence we can articulate a negative condition upon judgments of experimental incommensurability in the opaque regime: *whatever may be the intrinsic differences between the ingredients of two experimental practices*, we should not describe the situation as an experimental incommensurability of any kind when these ingredients are actually explainable with the help of a common theoretical framework.

Conclusion: Pickering's and Hacking's machinic-literal incommensurability can *not* be characterized on the basis of experimental features *only*. Their definition of machinic-literal incommensurability in terms of the disjunction of sets of machines or sets of measures is not complete: the very fact of experimental differences which, considered in themselves, might encourage a description in terms of an experimental *disjunction or an experimental ACM*, is *not sufficient* to impose an experimental incommensurability. At least an *additional* negative condition related to the situation of the high-level theories involved must be taken into account. We will now see that this additional negative condition is not yet a sufficient one: other conditions must also be satisfied.

6.2. To be or not to be competing scientific practices

6.2.1. The case of disjoint disciplines

What kind of scientific practices do satisfy the condition "no intersection between sets of machines and sets of measures" that defines Pickering's and Hacking's machinic-literal incommensurability? When we try to imagine concrete situations that would verify this description, one case immediately comes to mind: the case of distinct disciplines or distinct specialties.

Consider, for example, physics and experimental psychology. We can indeed describe their situation as a disjunction between the sets of instruments they use and the set of measures they obtain. So what? As long as physics and psychology are taken to be two distinct disciplines, that is, as long as they are considered to be two separate investigations of heterogeneous kinds of reality, this experimental ACM appears, at first sight at least, neither surprising nor interesting. The fact of an experimental ACM is precisely what is expected by everybody in the case of different disciplines. We are unable to identify a single conception of science that could be threatened or worried by a stress on the fact that distinct disciplines do not share common measures. This configuration appears totally innocuous. Hence the C condition is not satisfied for the

[20] See Sect. 6.3 for qualifications.

case that primarily comes to mind, namely distinct disciplines. Conclusion: the very fact of a machinic-literal ACM is not sufficient to generate a case of genuine incommensurability.

This points to the task of identifying what supplementary conditions must be met, in addition to the fact of an experimental ACM, to be entitled to describe the situation in terms of incommensurability. What, exactly, do we have to exclude, when we say that we must exclude the case of distinct disciplines and specialties? The answer I will now articulate is the following: negatively expressed, we must exclude the case of *non-competing* scientific practices; positively framed, we must include the *requirement of rivalry* as a *necessary condition* of experimental incommensurability. The first task is thus to discuss the notion of competition or rivalry.

6.2.2. What do we assume when we assume that two practices "compete"?

What do we endorse when we conceive two scientific practices as *competing* or as *rivals*? In detail this might of course vary depending on who is the "we". But at a general level, it is possible to elicit central commitments associated to the idea that two scientific practices compete.

When we say that two scientific practices compete, we assume, at the most fundamental level, that they are mutually exclusive in principle and hence that both should not be maintained in fact. This goes with the demand to choose either one or the other, or alternatively, to find a way to reconcile them in one way or another.

If we go a step further and ask what kind of assumption lies behind the idea that two scientific practices are mutually exclusive and cannot both be maintained, we find the commitment that the two researches under comparison should not be in conflict *because they study the same reality*: for example, the physical world for two competing physics, or a part of it for two competing physical specialities, or the human mind for two experimental psychologies. Let us speak of a rivalry related to the identity of the targeted object.[21] Since the two competing investigations are supposed to target the same object, they should ideally (*de jure*, in principle, at the ideal end of the research…) be associated with the *same measures* (and parenthetically also with the same theory). Hence if it is not the case in fact, if important tensions arise historically between the

[21] The rivalry ascribed to the identity of the targeted object, and its associated requirement of an identity of scientific content at all levels (measures, theories), is not the only kind of competition we can conceive. Two disciplines can also compete at the level of methodology. That will be the case, for example, if two physics or two distinct sciences (for instance, a physic and an experimental psychology) grow on the basis of methods that are differently valued and that are perceived as irreconcilable by practitioners. In such cases the competition is no longer ascribable to the assumption of an identity of the targeted domain of study (rivalry can indeed happen between disciplines that are supposed to study two distinct heterogeneous realities, as shows the example of physics and experimental psychology). The rivalry relates to a conception (or an ideal) of scientificity. Two communities endorse conflicting commitments about what a good science is, and hence conflicting commitments about the way a discipline should proceed in order to be a genuine science *rather than a false or pseudo-science*. In the present section I will set aside this rivalry of the methodological kind, since only the rivalry ascribed to the targeted domain is relevant with respect to the specific issue I want to discuss. We will meet the methodological kind of rivalry below, in Sect. 7.

measures (and between the theories as well) of disciplines that are thought to be in a relation of rivalry in the sense just specified, this will be perceived as a problem: practitioners will try to reconcile, *in one way or another*, the two sets of measures[22] (and the two different theories as well).

Admittedly, it is not easy to go further.[23] It is not easy to specify what an assumption of the kind "identity of the targeted object" means in detail.… It is not easy to explain why judgments of this type are or are not endorsed at a given stage of the history of science (why this particular mapping of disciplines holds).… And it is not easy to analyse how and why such judgments have sometimes evolved over time (creation/dissolution of disciplines and shift of the boundaries of existing disciplines). It is not easy, notably because we are dealing with deep commitments that indeed work in scientific practices but that are vague in nature and usually not made explicit by practitioners. But we can proceed without answering these questions, adopting once again the "opaque regime" of investigation.

As already indicated, in the opaque regime we take the actual commitments of a reference group (here commitments about rivalry/non-rivalry of disciplines) as a sociological fact and a primary datum. At this level, "non-competing" and "rival" are equated with "taken-as-non-competing" and "taken-as-rival" by a specified group of reference (the group of scientific practitioners or sometimes larger groups, especially when we deal with common classifications that cross our culture). At this level we take physics and experimental psychology as non-rival disciplines that deal with different domains of reality. Or we take what Pickering calls "the high energy physics of the 60s" and "the high energy physics of the 80s" as competing practices that target one and the same domain of reality.

6.2.3. The difference that makes a difference between competing and non-competing practices: rivalry as a necessary condition for incommensurability verdicts

There is an important difference, at the level of the literal common measure, between rival and non-rival practices. Rival scientific practices have, in principle, identical measures. Hence a fortiori they should have, in principle, a common measure in the literal sense. Hence they should manifest a literal common measure *in fact*. Now if this is not the case, it is astonishing and it raises a question. It is only under the presupposition that a literal absence of common measure should *in principle* be the case – a presupposition that is embedded in the assumption of rivalry – that the *fact* of a literal absence of common measure may appear epistemologically interesting and possibly harmful. If two scientific practices associated with distinct disciplines in charge of distinct domains of reality have literally no common measure, is not only admissible but

[22] This may be done in several ways. One may find a new unified theory able to explain both sets of measures.… Or one may disqualify one of the two sets as genuine experimental facts.… Or one may conclude that scientists were wrong in believing that the two sets were signs of the same kind of reality and had thus to be explained by one and the same discipline: one may ascribe one of the two sets to a distinct discipline and dissolve in this way the requirements associated to the condition of rivalry.

[23] I provide more elements in Soler (2006c).

moreover expected. In this case, the C condition is not satisfied at all. To the contrary, if two scientific practices dealing with the same domain are built on measurement results that have no intersection, that is a quite surprising fact. In this case, condition C could perhaps be satisfied. In short, it is only under the *expectation* of a literal common measure *in principle*, that the *actual fact* of an absence of literal common measure appears surprising, strange and possibly catastrophic.

The latter remarks have been applied, in the reflections just above, to the case of incommensurability *of the experimental kind*. But they actually carry a more general point. In fact, structurally similar considerations apply as well, *mutatis mutandis*, to incommensurability of the semantic kind between theories. It is only under the assumption that two theories compete, which in turn implies the expectation, in principle, of a semantic common measure between these theories, that the fact of a lack of semantic common measure appears astonishing and possibly catastrophic. This was recognised early in the debate on traditional incommensurability, notably by Feyerabend. As Martin Carrier wrote in a recent paper dealing with semantic incommensurability:

> The significance of incommensurability requires that there is some range of phenomena that the theories jointly address. (…) Darwin's theory of natural selection is not translatable into hydrodynamics; quantum mechanics cannot be rendered in the concepts of Zen. (…) Such cases need to be excluded. (…) In this vein, Feyerabend distinguishes between competing and independent theories and restricts incommensurability to concepts from theories of the former kind. (Carrier, 2001, p. 83)

I prefer to talk, negatively, of "non-competing" or "non-rival" theories rather than of "independent theories", since non-competing theories like distinct disciplines may entertain various relations, some of which do not correspond to a strict independence (e.g., experimental psychology could intend to take into account some physical features of human beings and would then not be totally independent of physics). But apart from terminological considerations, Carrier's reference to Feyerabend conveys, for the case of theories and semantic incommensurability, exactly the same point as that expressed above for the case of experimental practices and machinic-literal incommensurability.

True, we can understand what *may* lead to transgression of the above "prohibition" and to talk of an incommensurability in the case of *non-rival* practices. Kuhn himself, in some of his later articles, did this.[24] Spelling out the associated reasons will once again show why condition C is required.

[24] See Kuhn (1991). "Though I have in the past occasionally spoken of the incommensurability between theories *of* contemporary *scientific specialities*, I've only in the last few years begun to see its significance to the parallels between biological evolution and scientific development (…)". "After a revolution there are usually (perhaps always) more cognitive specialties or fields of knowledge than were before. Either a new branch has split off from the parent trunk (…). Or else a new specialty has been born at an area of apparent overlap between two pre-existing specialties (…). As times goes on, (…) one notices that the new shoot seldom or never gets assimilated to either of its parents. Instead, it becomes one more separate specialty (…). Over time a diagram of the evolution of scientific fields, specialties, and sub-specialties comes to look strikingly like a layman's diagram for a biological evolutionary tree. *Each of these fields has a distinct lexicon*, though the differences are local" (my emphases). Immediately following this passage, Kuhn talks about incommensurability's "role as an isolating mechanism". Although Kuhn does not use the expression of "non-rival practices", this is clearly what is involved when he speaks of separate scientific specialties.

If one may be tempted to talk of an incommensurability in the case of non-rival practices like physics and experimental psychology, it is for the following reasons:

1. The ACM condition is satisfied *twice*: first because there is a *machinic-literal* ACM between the experimental practices involved; and second because there is a *semantic* ACM between the associated high-level theories.
2. The ACM condition is satisfied to a degree that seems to exceed *by far* what can happen in the case of rival theories.
- At the level of measures: the machinic ends of two non-rival disciplines will be *more strictly disjoint* than the machinic ends of two rival disciplines. (In the first case we can meet the situation of a strict disjunction: not a single measure in common, the most extreme form of a literal ACM we can imagine.)
- At the level of theories: the theoretical taxonomies of two non-rival disciplines will be more ill-matched; they will possibly involve completely different concepts (as in the case of quantum physics and Zen) and they will possibly not have, even locally, a single specialized concept in common (described in terms of incommensurability, this situation could be labelled an *extreme* semantic incommensurability).

But besides these perfectly understandable reasons that might lead one to describe the situation in terms of an incommensurability, I maintain that the vocabulary of "incommensurability" should be avoided in such cases. Indeed, what is missing here is the satisfaction of the C condition. A strict machinic-literal ACM joined to an extreme semantic ACM does not have any catastrophic consequence in the case of *non-rival* theories, since in this case, at the *experimental* level there is *no* in-principle expectation to the effect that the measurements *should* be the same, and at the *theoretical* level there is no in-principle demand of conceptual homogeneity.

In conclusion, the condition of rivalry is a necessary condition of an incommensurability verdict, be it of the experimental or of the semantic type. This was recognized early in the discussion of semantic incommensurability. Now it appears that the same holds for experimental incommensurability. This condition is not discussed or even mentioned by Pickering and Hacking.

6.2.4. Qualifications: rivalry in the past and rivalry in the present

Actually, the condition of an actual rivalry, as it has just been described, is not the last word. We must take into account not only rivalry in the present, but also rivalry in the past. In this section, I will argue that non-competing disciplines in the present can be said to be incommensurable if there is a sense in which they have been rivals in the past.

Let us see how we can make sense of the idea of two scientific practices that, whereas presently viewed as not rival, have nevertheless been considered as rival *in the past*. Imagine for instance two contemporary experimental practices EP1 and EP2, originally viewed (say at time t) as two rival practices coordinated with a unique discipline (say physics), originally associated with two disjoint sets of instrumentaria and measures, and originally coordinated with two apparently irreconcilable theories. Imagine now that the conflict between EP1(t) and EP2(t) has been, historically, solved in the following way: at a subsequent stage t' of the history of science, EP1(t'), the descendant of EP1(t), is still viewed as a physics, but EP2(t'), the descendant of EP2(t), is now associated with another different discipline, one responsible for a heterogeneous domain of reality (say psychology, but more radically we can imagine a new discipline unknown to our present classifications). In the most extreme version of such a configuration, EP1(t') and EP2(t') are based on *strictly* disjoint sets of instrumentaria and measures, and they are explained by two *completely* heterogeneous frameworks: in other words there is, between EP1(t') and EP2(t'), both an strict ACM at the literal level, and an extreme ACM at the semantic level.

This indeed exemplifies a situation in which EP1 and EP2 do not compete in the present (at t') but were in competition in the past (their ancestors were rivals at t). If we only consider what holds at t', we will not expect, *in principle*, any experimental and semantic common measure between EP1 and EP2. Hence when we do not find any *in fact*, we are not surprised. We will not consider this fact as an instructive epistemological situation, and we will not associate with it any potentially damaging consequence. But if we turn to the past history of EP1 and EP2, we will discover that past scientists strongly expected an experimental and semantic common measure between EP1 and EP2. From their point of view, the fact of an experimental and a semantic ACM between EP1 and EP2 is unexpected, remarkable, strange and potentially damaging: what had to be commensurable is in fact not.

What potentially damaging consequences are involved here? If we deploy the most extreme possible consequence of such rival - > non-rival transitions, we can conceive a counterfactual science that, starting from the same initial state as ours, would, after a sufficiently long time, map the world in a completely different manner as we do, by means of a heterogeneous, non-homologous cartography of disciplines. Pushing things to the extreme limit, what lurks behind such rival - > non-rival transitions is the idea of a counterfactual science that would have nothing in common with ours except a certain general conception of scientificity: a science that would share with ours no discipline, no instrumental device, no measure, no theoretical assumption. *If* such an extreme contrafactual alternative were plausible, it would be damaging for inevitabilists and many realists.[25] The situation would warrant, if not require, the lexicon of the incommensurable. We would have a radical incommensurability between our science and the alternative other one. It would be a global incommensurability between two stages of scientific developments considered as wholes.

Now of course, the plausibility of such an alternative science is questionable. Inevitabilists, many realists and most scientists will argue that such a configuration, if it is conceivable, is not plausible at all: they will see it as a sceptical philosophical fiction that is not informative at all about our real science. Contingentists like Pickering will argue the opposite. Admittedly, the issue is not easy to decide. But with respect to our decisions about incommensurability verdicts, what matters is not who is right and who

[25] It would be damaging at least for the "correspondentist" realist, who assumes that essential parts of our scientific theories correspond to the studied reality (who assumes, for example, that there is a physics and a biology because there indeed exists two different kinds or levels of reality, the physical and the biological).

is wrong: it is the very existence of a conflict that has important implications for the two opposed camps. Such conflict loads the fiction of a radically heterogeneous science with a catastrophic *potential*.

Even if we refrain from extrapolating, from the rival - > non-rival *local* transition described above, to its most extreme version of an alternative science *globally* different from ours at all levels, the *local* transition *alone* is, already in itself, not deprived of a catastrophic potential. Such a local transition is not just a fiction created by the philosopher for the sake of the argument. We could find historical examples of such a configuration. True, historical case studies might always be dismissed as unfaithful or irrelevant as exemplifications of this or that philosophical characterization. But the status of the rival - > non-rival *local* transition is nevertheless different, closer to the "real", than the status of the radical global extrapolation, which is the status of a modality. In any case, the local case raises the question of the origin, and of the possible evolution, of our fundamental assumptions concerning which phenomena are of the same kind and should be explained by one and the same discipline, which phenomena are of different kinds, and what kinds of kind are involved in each case. As soon as there is a doubt about the stability of such assumptions, there is a source of conflict between inevitabilists who believe in a sufficient stability at this level, and contingentists who think that a deep instability is plausible. This conflict loads the rival - > non-rival local transition with a catastrophic potential. It possesses exactly the same structure as the global extreme case, even if its implications will probably appear weaker to most participants in the debate, and even if the discussions will probably appeal to partially different argumentative pieces.

All in all, I think that in cases structurally similar to the rival - > non-rival local transitions described above, we are entitled, if not required, to characterize the strict machinic-literal ACM between EP1(t') and EP2(t') as a machinic-literal *incommensurability*, and the extreme semantic ACM between EP1(t') and EP2(t') as a semantic *incommensurability*. The fact that EP1 and EP2 are historically derived from one and the same discipline (physics in the example), or in other words, the fact that EP1 and EP2 have had *in the past*, *in principle*, a literal and semantic common measure, makes non-trivial, epistemologically relevant and interesting, the issue of their actual comparison at the literal and at the semantic levels, despite the fact that EP1 and EP2 are, in the present, seen as two non-rival disciplines for which no living practitioner would require a literal and semantic common measure.

6.3. A decisive, bottom-up influence exerted by experimental factors?

Let us sum up our conclusions. In order to be a candidate for characterization in terms of experimental incommensurability, two experimental practices must satisfy at least two conditions:

(a) Not to be actually explained by a unitary high-level theory or by two different but addable theories

(b) To be competing practices (i.e., to be viewed as targeting the same object and hence as *in principle* "additive" in terms of experimental means and results), or to have been competing practices in the past

These conditions put the stress on factors that are apparently *not* characteristic of the *experimental* level proper. The first one explicitly concerns what holds at the *theoretical* level. The relation of the second one to specific features of available experimental practices would require discussion. But in any case, at first sight there is a doubt: does the property of rivalry involved in (b) (the practitioners' global conception of the different kinds of reality) strongly depend on specific features of available experimental practices?

On the basis of such remarks we might be tempted to ask: is there ultimately any harmful consequence of the possibility of experimental disparities *considered in themselves independently of the associated high-level theories and of general commitments about the kinds of reality that constitute the world*? Is not the possibility of an experimental ACM, although called "incommensurability" by Pickering and Hacking, in the end completely innocuous? Are not experimental disparities unimportant, inefficient, essentially controlled by what happens at other levels and notably at the theoretical one? In this section, I will lay out what should be assumed about experimental practices in order to give a negative answer to these questions.

In order to attach harmful epistemological consequences to the fact of an experimental ACM, we must ascribe a causal power to experimental differences – or, if "causal" sounds too strong, at least a capacity to influence the historical trajectory in decisive ways. In that vein one would argue, for instance, that it is *because* the instrumentaria and the associated experimental marks were mutually exclusive at a given time that physicists did not manage, at a subsequent time, to unify the two associated sets of experimental marks under a single high-level theory. Or one would claim that it is *because* the instrumentaria and the associated experimental marks were mutually exclusive that this or that field has subsequently split into two different fields.

If such bottom-up influence could be exhibited,[26] would we be entitled to talk about an experimental *incommensurability*? That is: would our C condition be satisfied? What conception of science would be potentially threatened? The answer I will articulate is the following. Inevitabilist and many realist interpretations of scientific advance would be potentially threatened, if contingentists and anti-realists were to succeed in arguing two related points:

(a) Contingent features of experimental practices – features that could have been completely different in other possible roads of scientific investigation – are responsible for essential aspects of our science
(b) Disparate experimental practices, for example practices based on disjoint instrumentaria, would have led to a science irreducibly different from ours

[26] Nickles seems to admit something of this kind, when he writes, in his contribution to the present volume, that "theoretical-representational ruptures" "are not always the cause of" "disruptions of practice". "In fact, in some cases the direction of the causal arrow is reversed".

When we analyse this answer, we can distinguish three points:

(1) First, the bottom-up influence at issue must be interpreted as due to *contingent* features of experimental practices.

Indeed, if the influencing experimental features were not contingent features, for example if they were induced by the physical reality under study, realists and inevitabilists would not be worried at all. To assume that the influencing experimental features are contingent amounts to assuming that the reality under study could have been investigated by very different experimental means – up to the extreme case of completely disjoint instrumentaria. Assumptions of this kind are commonly endorsed for experimental *means* alone. Framed in terms of means *only*, such a claim would not upset any realist or inevitabilist, since both can still maintain that different means can lead to the same results. Hence the need for supplementary conditions (cf. point (3) just below).

(2) Second, the bottom-up influence at issue must be interpreted as due to *intrinsic* features of experimental practices in the sense of "sufficiently independent of high-level theories".

Otherwise we would not be dealing with an *experimental* influence *proper* and we would not consider it a candidate for *experimental* incommensurability. As any philosopher of science knows, it is not an unproblematic task to exhibit "intrinsic features" of experimental practices. But here I just want to indicate what should be assumed in order to transform an experimental ACM into a case of genuine experimental incommensurability.

(3) Third, the bottom-up influence at issue must be interpreted as a *decisive* influence.

What does this mean? The causal power ascribed to the experimental practices must influence the history of science *in a way that makes a non-reducible difference*. That is: two *irreconcilable* sciences could be produced on the basis of two disparate experimental practices, *because* of the experimental disparities involved.

On the basis of these three points, we can conceive of an alternative science that would, *due to differences at the experimental level*, be different at all the other levels. We can again generate the extreme idea of a counterfactual science that would have no common measure with ours at the fundamental level of basic disciplines and specialties: a science that would map the targeted object completely differently. If such possibilities were plausible, the fact of an experimental ACM could legitimately be described as an experimental *incommensurability*.

Now, as in the last section we can ask: are such possibilities *plausible*? It is clearly very difficult, and may lie beyond our capacities, to adjudicate between the stronger sense and other possible weaker senses of the experimental influencing power under scrutiny. True, the contingentists, the holists and all those who favour a path-dependant model of scientific progress, have plausible arguments in favour of the idea that, *had* the instrumentaria been different, then our science would have been *different* at all levels. But this is a weaker claim than the one involved in the three points above, since "*different* sciences" does not necessarily imply *incompatible* sciences. And it is easy for inevitabilists or realists to reply that nothing proves that this alternative science would be irreconcilable with ours. Inevitabilists and many realists are, as for them, prone to a reading of the history of science that has the effect of nullifying the damaging potential of the causal power ascribed to the experimental disjunction. According to their interpretation, the contingent elements of the experimental practices have at most the power to determine some minor details of the historical trajectory. They do not hinder, but only postpone, what had inevitably to occur in terms of the experimental data and of the theories of these data, in virtue of the very nature of the world (and possibly also of the very nature of scientific methods).

Hence the issue is not easy to decide.[27] But from a logical point of view, *if* we could admit a decisive bottom-up influencing power, from the experimental practices (for instance, from a machinic-literal disjunction) to the high-level theories (for instance, toward the impossibility of a unifying theory) or to the disciplinary classifications, *then* the condition C would be sufficiently satisfied to allow the interpretation of an experimental ACM in terms of an experimental *incommensurability*. In this case it would remain to analyse, in the transparent regime, the ways in which the incommensurability of *experimental practices* may be related to high-level theories, in particular the ways in which an experimental incommensurability may create some semantic and/or methodological incommensurabilities at the theoretical level.

7. A CENTRAL EXAMPLE OF PICKERING'S EXPERIMENTAL INCOMMENSURABILITY, REDESCRIBED AS A SPECIAL CASE OF METHODOLOGICAL INCOMMENSURABILITY

In the last section, I examined the idea of experimental incommensurability in the opaque regime. A full investigation of the problem in the transparent regime would require opening the box of experimental practices in order to characterize the kinds of experimental differences that may lie behind an experimental incommensurability. In this section, I will (partially) open the box of a particular historical example studied by Pickering under the heading of incommensurability: the "discovery of the weak neutral current".

7.1. Re-analysing one of Pickering's central examples: A methodological incommensurability of experimental procedures

In *Constructing Quarks*, Pickering explicitly introduces the episode of the discovery of the weak neutral current as a case of "local incommensurability" (Pickering, 1984, p. 409) arising at the experimental level.[28] His description of the corresponding situation is, briefly stated, the following. Depending on the local "interpretative practices" that experimenters rely on to interpret the pictures of neutrino experiments obtained with bubble-chambers, physicists are led either to deny or to admit the existence of the weak neutral current.

[27] For a deeper analysis of the difficulties related to the contingentist/inevitabilist issue, see Soler (2007).

[28] "Incommensurability is visible in the history of HEP at the local level of the performance and interpretation of individual experiments. The neutral current discovery can stand as an example of this" (Pickering, 1984, p. 409).

The example involves two (in this case successive) *rival* particle physics coordinated with two *rival* scientific practices. These incommensurable practices are rivals since practitioners assume that they target the same domain – the microscopic constitution of matter. Since they are rivals they should, in principle, have a common measure, literally as well as semantically: actually they should *ideally* be associated with the *same* (or at least reconcilable) measures, with the *same* (or at least reconcilable) semantic frameworks and with the *same* (or at least reconcilable) theories. But this is not the case *in fact*.

What is the actual situation? The two particle physics are semantically commensurable at the theoretical level (at least concerning the weak neutral current). They use the same kinds of instrumental devices (bubble chambers). And they are built on a common set of experimental marks (a common set of photographic plates). Nevertheless, they assume mutually contradictory claims about the existence of a particular phenomenon, namely the weak neutral current. And Pickering relates this theoretical difference to a difference at the experimental level. More precisely, he points to a particular variable ingredient of the experimental practices involved, in the 1960s on the one hand and in the 1970s on the other hand, in neutrino experiments interpreted with visual detectors[29]: Namely, the fact that experimenters of the 1960s and those of the 1970s relied, very locally (for experiments with neutrino beams and visual detectors), on two different "interpretative procedures" in order to convert the experimental marks (what is visible on the photographic plates) into phenomena (existence/inexistence of the weak neutral current).

Basically, the interpretative practices at stake involved energy "cuts" that led physicists to disregard, on the pictures obtained with bubble-chambers, all events having a total energy smaller than a certain threshold. Historically, two different procedures of energy cut were applied to bubble-chamber pictures, one in the 1960s when the weak neutral current was assumed not to exist, and the other in the 1970s when its existence became a crucial issue for physicists and when the phenomenon was finally considered to have been discovered in 1974. The corresponding scientific practices can be viewed as mutually exclusive in the sense that physicists cannot be committed to both of them at the same time. Pickering describes them as incommensurable.

Let us pause here and make two comments.

First, in this example, the disjoint local interpretative procedures involved are not in themselves material objects. Hence the example clearly shows that *machines*, *qua* material objects, are not, in Pickering's writings, the only possible source of incommensurability at the level of experimental practices. This could lead us to question the appropriateness of the "machinic" idiom for cases of that kind.[30]

[29] Actually a similar argument applies, according to Pickering, to the experiments that use electronic detectors, but I will here leave them aside.

[30] Pickering's own position about this appropriateness is not so obvious. As already stressed, Pickering does not use the expression "machinic incommensurability" in Pickering (1984) when he first analysed the example of the weak neutral current at length. When this terminology is introduced in *The Mangle of Practice*, the illustration he articulates is a global one: it involves the old HEP of the 1960s and the new HEP of the 1980s *taken as two wholes*, not just the local ingredients related to the weak neutral current that constitute a part of each of them. Now one thing is at least sure: Pickering takes the example of the weak neutral current as an example of experimental incommensurability, and in the analysis of this example, he insists on the importance of material aspects of experimental practices.

Second, the example does not appear as a good instantiation of a case of *disjunction* of experimental practices: a lot is shared between the two experimental practices under scrutiny taken as two wholes, including at the level of machines (bubble chambers in both cases) and at the level of measures (photographs issued from bubble chambers). True, we could also list other important differences between these experimental practices in addition to the difference already mentioned concerning energy cuts: for example the degree of mastering of neutrino experiments and the related quantity of reliable data about them; or again, very important, the size of the bubble chambers involved, which have crucial repercussions for the type of information available on the pictures.… But it does not seem appropriate to speak of a disjunction between the two experimental practices. Or should we say that there is a *partial* disjunction between them? Or a local disjunction concerning only particular, well-delimited ingredients of them? If yes, we should at the same time recognize that the experimental disjunction may be more or less extended. But we can as well see the example as a kind of experimental incommensurability different from the "machinic" kind.

This being said, let us see in what sense our ACM and C conditions are satisfied in such a situation.

The ACM condition is fulfilled in the following sense: The two local interpretative procedures of energy cut amount to two different norms governing the reading of photographic plates. The lack of common measure corresponds to the fact that *very locally*, certain *incompatible standards are adopted* (here two interpretative procedures locally implemented to turn the machinic ends into phenomena).

What about the C condition? What are the catastrophic consequences of the existence of two locally different interpretative standards? Pickering's answer is something like the following.

The local interpretative procedures (LIPs) of energy cut are constitutive of what is taken as the existing phenomena. They are thus constitutive of the ontology (since the phenomena involved here are highly theoretical entities such as quarks or weak neutral currents). In our particular case, to accept one or the other of these two interpretative standards leads to the acceptance of mutually contradictory statements concerning the existence of a certain phenomenon (the neutral current NC).[31] In brief, we have the following scheme:

If LIP1, then no NC; and if LIP2 (all other things being equal), then NC.

Furthermore, the choice between LIP1 and LIP2 is "an unforced and irreducible choice" (Pickering, 1984, p. 405). It is so in the sense that it is not determined by some necessarily compelling reasons at a given stage of scientific research. In particular, neither "nature" nor anything like "*the* scientific method" can impose the choice. More concretely, the machinic ends do not dictate the choice. On the basis of the same performed experiment and the same resulting photographic plates, physicists of the 1970s

[31] "To choose between the theories of the different eras required a simultaneous choice between the different interpretative practices in neutrino physics which determined the existence or non-existence of the neutral current. The later choice (…) cannot be explained by the comparison between predictions and data which were its consequences" (Pickering, 1984, p. 409). Just after Pickering speaks of phenomena "brought into existence by appropriate 'tuning' of interpretative practices".

could have conserved LIP1 instead of changing it for LIP2, and thus could have lived in a world in which weak neutral currents do not exist.

We are approaching a contingentist position often classified as "relativist". A relativism of this kind appears damaging to many philosophers of science, and especially to inevitabilists and most realists. Hence we are entitled, according to our preliminary decisions, to conclude that the C condition is satisfied, and to speak of *incommensurability*.

What kind of incommensurability? As far as it holds between ingredients of experimental practices, we can describe it as an *experimental* incommensurability. But is this experimental incommensurability a new *kind* of incommensurability, different from the two traditional semantic and methodological kinds? No. It is a special case of methodological incommensurability. It is a *special case* insofar as it concerns a particular constituent of experimental practices (namely the LIPs used by physicists working in neutrino experiments in the 1960s and the 1970s in order to convert the tracks on photographic plates issued from bubble chambers into existing particles). But it is a case *of methodological* incommensurability, since it is a case in which two incompatible standards are at work. In brief, Pickering's example is a case of methodological incommensurability that is not classified as such.

7.2. Generalisation: An experimental, methodological incommensurability

With the example of the weak neutral current we found, at the level of *experimental* practices, an analogue of the methodological incommensurability traditionally applied to the norms governing the elaboration of high-level theories on the basis of unproblematic experimental data. It seems that we can, starting from the example, generalize the result.

- We can generalize it by enlarging the phenomenal scope:
  The example deals with a limited phenomenon, the neutral current, but we can imagine that the same scheme underlies a configuration in which two contemporary or successive scientific communities accept the existence of large disjoint classes of incompatible phenomena.
- We can generalize the result by diversifying the ingredients of the scientific practices acting as standards:
  The example refers to two local interpretative practices and their methodological incommensurability, but we can imagine that many other elements or combinations of elements involved in the experimental practices may act as incompatible standards for different groups. This can be the case, for instance, with measurement instruments. Actually, any ingredient of scientific practices can be turned into a standard. Hence we can imagine a methodological incommensurability of experimental means in general.

From there, we can derive the most general definition of the methodological incommensurability of experimental practices.

It can be defined as the recourse to different experimental norms leading to the adoption of irreconcilable experimental data, the decision to accept one or the other of these standards being "an unforced and irreducible choice" (Pickering, 1984, p. 405).

The ACM condition is satisfied in the following sense:

The experimental means deemed relevant for… or reliable for… or better for... the measurement of this or that theoretical variable or ensemble of variables.… or the investigation of a certain kind of problem... or the like… fall into two (at least partially) disjoint ensembles. In other words, there is an incompatibility between judgments of relevance-reliability-superiority about certain experimental means.

The C condition is satisfied in the following sense:

There are *irreducible* disagreements about the *relevance* or the *meaningfulness* of the same machinic ends, with the consequence that the history of science is fundamentally contingent, that it could have been different in essential respects, including with respect to what is identified as robust experimental data. The methodological incommensurability of experimental means is damaging for every philosopher who cherishes the idea that there are sufficiently strong (if not necessary) constraints upon scientific investigation: constraints that act (or should act on good scientists) in a compelling way – at least in some crucial stages of the research, once a problem has been "sufficiently" investigated (these constraints being most of the time related in one way or another to the object under investigation and its "pressure"). It is in particular damaging for inevitabilists and for most realists. The methodological absence of common measure can thus be interpreted as a methodological *incommensurability* between some experimental means.

If there is a novelty here, it concerns not the kind of incommensurability, but the answer to the question: "the incommensurability *of what*?"

The answer is not only, as traditionally: of the norms that govern theoretical assessment on the basis of a set of unproblematic measurement results. It is now enlarged to any ingredient of the experimental practices *constitutive of what plays the role of experimental data*. It seems that methodological incommensurability may apply, potentially, to any experimental means that works as a norm and contributes to constitute experimental data. I think that people who are interested in the study of experimental practices would gain by recognising as such cases of methodological incommensurability.[32] This would allow them to profit from the conceptual tools and the analyses that have been elaborated in traditional discussions of the incommensurability problem.

8. THE PART AND THE WHOLE PROBLEM: SKETCHING THE IDEA OF AN INCOMMENSURABILITY *BETWEEN SYMBIOTIC PACKAGES OF HETEROGENEOUS ELEMENTS*

In Sect. 7, I applied incommensurability to the relation between two very circumscribed ingredients of particle physics in the 1960s and the 1970s, namely two local interpretative experimental procedures. From this characterization focused on a single well-delimited ingredient of experimental practices as the answer to the question "the incommensurability *of what*?", one might understand that, historically, the "unforced and irreducible" choice of practitioners just took place between these two ingredients

[32] It seems to me that at least part of Nickles' "recognition problem" – conflicts between practitioners about what is relevant to their practice and what is not – is a case of methodological incommensurability (or value incommensurability).

and only them: or in other words, that practitioners had to choose *just* between the two experimental interpretative procedures under scrutiny (and consequently between the negation/affirmation of weak neutral currents) *all other things being equal*. Moreover, the reference to an "unforced and irreducible choice" quite irresistibly suggests that practitioners had no compelling reasons to choose one option rather than the other. It suggests that practitioners were "free", in the sense that no pressure at all was exerted on them or at least in the sense that no universal reason made them favour one or the other of the two options.

But reading Pickering more carefully, we can at least suspect that such an understanding may not be faithful to his conception. More generally, we can suspect that it is perhaps not the best way to characterize the situation: that it is perhaps only a partial, and in certain respects a spurious characterization. To investigate this point would require another paper, but in this last section, I want at least to sketch the content of such doubts and the principle of what could be a potentially more complete and exact alternative explanation. This will indicate a possible direction for future research, with the hope that other philosophers will take such a road.

The doubts at stake have many possible origins. For example, Pickering's account of the evolution of the particle physics from the 1960s to the 1970s stresses, besides the two particular local interpretative procedures of bubble-chamber pictures, many other important changes; and this may lead us to question the legitimacy of a description that presents the situation of practitioners as a choice between just two experimental procedures *all other things being the same*. More fundamentally, Pickering offers an explanation of the *reasons why* the new physics has been adopted to the detriment of the old one, and this explanation, first does not lead to conceive practitioners as free in the sense of "no pressure at all"; and second *essentially* involves *other* ingredients than the two local interpretative experimental procedures. I said "essentially", since the explanation at issue is holistic in nature. The same kind of holistic explanation is endorsed by (Hacking, 1992). It is granting an explanation of this kind that we can be tempted to substitute, to the local explanation presented in Sect. 7, a more global explanation. I will label the corresponding scheme of scientific development "symbiotic", adopting an adjective often used in Pickering's and Hacking's writings. In the next section, I will briefly present the symbiotic scheme of scientific advance. Then I will sketch the shifts this scheme could induce with respect to the questions "the incommensurability *of what*?" and "an incommensurability *of what kind*?"

8.1. *The symbiotic model of scientific advance*

According to the symbiotic model there are, in the history of science, no privileged types of factors, pre-determined and identifiable a priori, that are (or should be) the motor of scientific change. What determines the decision of scientists lies in the *global balances* achieved. At some points in the history of science, the reciprocal reinforcement and the mutual stabilization of multiple factors of different types (material, behavioural, propositional…) constitutes a "scientific symbiosis": a complex package that is, in a sense, "self-consistent", "self-contained", "self-referential" (Pickering, 1984, p. 411) and "self-vindicating" (Hacking, 1992). Scientists' decisions depend on the *degree of perfection of the available scientific symbioses*.[33]

As a matter of fact, the nature of the constitutive elements of scientific symbioses is a topic of discussion between the authors who favour the symbiotic model. Pickering's and Hacking's positions on this point seem themselves not consensual.[34] But beyond these divergences, the general structure of the common underlying scheme is clear. It is a holistic and an interactionnist scheme, a scheme of the co-stabilization based on an enlarged idea of coherence (enlarged to take into account non-propositional elements[35]).

Although more work is still needed in order to specify the picture, I think this is a promising way to model scientific development. Now, if such a model is taken for granted, the two terms between which practitioners had to choose in the example of the weak neutral current discovery were not simply two local experimental procedures. If the ultimate justification of actual scientific options lies in the relative quality of the available alternative symbioses, the two terms between which practitioners had to choose were, instead, two complex symbiotic packages, each of which includes one of the two experimental procedures as a constitutive element (and also other differences). Practitioners chose the second holistic package (including the local interpretative procedure of the 1970s and the existence of weak neutral currents) because it offered a better symbiosis than the first (including the local interpretative procedure of the 1960s and the no-neutral-current assumption).

But if this is correct, shouldn't we revise the characterization given in Sect. 7, when we spoke of an ACM and an incommensurability *between two local experimental standards* like two local procedures of energy cut?

8.2. An ACM between scientific symbioses?

If the relevant units of analysis indeed are, for living practitioners, whole packages rather than any local element, should not the philosopher of science himself shift his scale of analysis, and try to characterize the lack of common measure (and the possible associated incommensurability) *globally*, as the emergent result of intertwined differences between two scientific symbioses rather than as local differences between two particular ingredients of these symbioses? In that perspective we would have an ACM

[33] "The communal decision to accept one set of interpretative procedures in the 1960s and another in the 1970s can best be understood in terms of the symbiosis of theoretical and experimental practice" (Pickering, 1984, p. 188). "When the laboratory sciences are practicable at all, they tend to produce a sort of self-vindicating structure that keeps them stable". "As a laboratory science matures, it develops a body of types of theories and types of apparatus and types of analysis that are mutually adjusted to each other. They become (…) 'a closed system' that is essentially irrefutable. They are self-vindicating in the sense that any test of theory is against apparatus that has evolved in conjunction with it – and in conjunction with modes of data analysis. Conversely, the criteria for the working of the apparatus and for the correctness of analyses is precisely the fit with theory" (Hacking, 1992, p. 30).

[34] See Hacking (1992).

[35] "A coherence theory of thought, action, material and marks" (Hacking, 1992, p. 58). See also Bitbol (2004).

between scientific symbioses (and hence possibly an incommensurability *of scientific symbioses*). The answer to the question: "the incommensurability *of what*?" would be: of *scientific symbioses*.

One may protest at this point that there is a sense in which we can speak, as we did in Sect. 7, of a local ACM of the methodological type. And indeed there is one: nothing prevents us from using the expression "a local ACM of the methodological type" to refer to situations in which two incompatible experimental norms are applied locally. However, if the value or significance of a component like a local experimental procedure is not given per se, if it essentially depends on its relation of mutual support with many other components of scientific practices, one may reply that whatever expression is used to name situations in which two incompatible experimental norms are applied locally, and despite the fact that descriptions like "two incompatible experimental norms are applied locally" are perfectly admissible descriptions of the historical situation, such a characterization cannot be the last word. One may argue that although it is not false, this characterization proceeds from an incomplete and truncated analysis of the historical situation. One may argue that it is therefore, if taken as the last word, a misleading characterization. This would be because the two experimental procedures have been artificially extracted from their original packages, while their physical significance, as well as the comparative judgments of their relative merits, essentially involve other elements and ultimately require one to take into account large packages that work holistically or "symbiotically". In that perspective, one could think relevant to introduce the idea of (global) ACM *between scientific symbioses* in addition to the (local) ACM between particular experimental ingredients.

But in that case, is it still appropriate to talk about an *experimental* ACM – and possibly an *experimental* incommensurability? If we must take into account large packages including multiple heterogeneous ingredients, aren't we going to be led to take into account ingredients of the theoretical sphere also? It depends, of course, on the delimitation of the relevant symbiotic unit in each particular historical situation. With respect to this point, it has to be stressed that the task of identifying (and thus of delimiting) the relevant symbiotic units is a difficult task. This is because nothing determines in advance which elements are constitutive, and because we must succeed in showing, in each case, that a certain network of relations actually does play the role of a mutual support and grants to certain items the status of constitutive elements of a well-delimitated symbiotic unit that is, in a sense, self-consistent and self-vindicating. It is an intrinsic difficulty of the symbiotic model.[36]

[36] I want to stress that in Pickering's writings, there are important fluctuations about the packages that are said incommensurable. They tend to change from place to place (and they are not restricted to the level of instrumentaria). If we try to classify thematically the occurrences from the narrower one to the most encompassing one, we find that what is said incommensurable corresponds to:

- Sometimes more or less precisely defined ingredients or combinations of ingredients of experimental practices (the devices being a particular case)
- Sometimes two "experimental practices" treated as unities considered in isolation with respect to theoretical practices
- Sometimes two "scientific practices" equated to two symbiotic packages "theoretical practice + experimental practice"

Now with respect to this question of delimitation, we can stress that in the example of the discovery of weak neutral currents, the relevant scientific symbioses involved are not restricted to elements of *experimental* practices.[37] Admittedly there could be other cases in which there are. But here, following Pickering's own analysis, we can *not* understand why scientists have implemented the two energy cut procedures one after the other, and why they finally chose the second one, if we limit ourselves to considering only factors pertaining to experimental practices taken in isolation from higher-level theories. The reliability granted to each of the two local interpretative procedures is constitutively linked to some highly theoretical assumptions (and reciprocally). The physical significance and value of the two procedures of energy cut is determined holistically, depending on the experimental *and theoretical* context involved – and this context has been notably modified *both* at the experimental *and theoretical* levels from the 1960s to the 1970s. Hence one will never be able to understand why the new experimental standard replaced the old one if one only considers, in isolation, just the two local experimental procedures of energy cut or even just the experimental practices in isolation. Considering them in isolation, there seems to be no reason to choose one or the other. Reasons can only appear to the person who zooms back and takes into account broader packages: reasons appear when one considers the two symbiotic units that establish a relation of mutual support between experimental elements and *theoretical elements*.

8.3. An ACM of what kind *between scientific symbioses?*

If we admit the idea of a global ACM between scientific symbioses, the next question is: an ACM (and a possible incommensurability) *of what kind* between scientific symbioses?

The answer will of course depend, in detail, on the nature of relevant symbioses that the analysis will have delimitated. But at the general level, we anticipate that in most cases, the global ACM at issue will not be reducible to "pure" traditional kinds. In most cases, we will find many heterogeneous kinds of local differences – for instance material ones, behavioural ones, semantic ones, methodological ones.… These differences could appear insignificant considered one by one, but *taken altogether and intertwined one with the other*, they will constitute two disparate units that will appear deeply different as wholes: as wholes they will resist point-by-point comparison but they will nevertheless, considered one independently of the other, appear "self-consistent" and "self-vindicating". The different ingredients of each package will nicely hang together, and each package, considered in itself independently of the other, will appear convincing.

[37] Admittedly, there could be other historical cases in which the experimental practice under scrutiny could be more autonomous. In this respect, see Galison (1997) and his idea of "intercalated periodization", or in other words the idea that theoretical and experimental practices might be partly autonomous subcultures having their own tempo and dynamics of change.

Since all the ingredients contribute to the robustness of the whole, all of them should be taken into account for the analysis of the ACM between the two wholes – including the material ones. Pickering especially insists on material elements (in the broad sense of "material") because he thinks they have been damagingly neglected (notably in traditional approaches to incommensurability). True, a "material" difference like the fact that experimenters of the 1970s worked with bubble chambers much bigger than physicists of the 1960s, or any other material local difference between two scientific practices, appears insignificant considered in isolation. But intertwined with many other elements – for example this renormalisable gauge-theory of electroweak interactions (not available in the 1960s), and this procedure of energy cut for the interpretation of bubble-chamber pictures (available in the 1960s but not applied), and so on – it becomes significant as one essential ingredient that contributes to the robustness of the whole and to the constitution of what is taken as a scientific result. Pickering's point is that one should not consider such "material" elements as unimportant with respect to what finally emerges as a scientific (experimental or theoretical) result. These elements must be included in the picture as *constitutive* elements, not just as contingent historical conditions of results that had to be obtained given what the world is and the questions scientists asked (e.g., do weak neutral currents exist?).

In such a conception, the global ACM would be the totalizing effect of a complex mixture of heterogeneous kinds of local differences: the result of the combination and the articulation of a multiplicity of local differences. The global ACM could be analysable into a multiplicity of local heterogeneous differences (possibly described as local ACMs). But the practitioners' positions and choices could not be understood at the local level: they could only be explained by taking into account the articulation of the different elements that nicely hang together in each package.

Let us grant that it is an adequate way of describing at least some episodes of the history of science. Could such a global ACM at least sometimes correspond to a global *incommensurability*?[38] What could be the damaging consequences, and for which conception of science?

8.4. From the idea of a global ACM to the idea of a global incommensurability

The answer to this question depends on the way one analyses scientific symbioses and thus understands their compelling power. In Pickering's explanation as I reconstructed it in Sect. 8.1, the "ultimate reasons" for scientific choices are related to the "quality of symbiosis" factor. What, exactly, does this mean?

It means at least that "nature", represented by the "machinic ends", does not dictate the choices, since scientists' decisions cannot be explained in terms of whatever particular element constitutive of each symbiotic package (be it the "machinic ends",

[38] Pickering uses the expression "global incommensurability", but not for the case of the weak neutral current. He speaks of a "global incommensurability" between the particle physics of the 1960s and that of the 1970s. This incommensurability encompasses several dimensions, notably the experimental, material one, but also the theoretical one.

or the local experimental procedures, or any other ingredient). Now does this position imply that choices between such packages lack any reasonable ground or any good reasons? This is the issue that should be investigated in order to decide if a global ACM can be seen (at least in some cases) as a global incommensurability.

The legitimacy of characterizing *as incommensurabilities* the missing common measures found between two symbiotic units, will depend on the way one analyses the ultimate factor "better or worse quality of symbiosis", on the three levels of:

(a) The types of ingredients that one is ready to deem constitutive of a scientific symbiosis (nature of the elements and of the relations of mutual support).
(b) The relative weight that one is ready to grant to the different types of ingredients.
(c) The degree and uniformity of the compelling power ascribed to the symbioses as resulting units: the force and the uniformity with which one admits that the "quality of symbiosis" factor governs scientists decisions.

According to Pickering, the constitutive elements of the symbioses can, in principle, be of any nature. They can notably include (as they do in general) some components that are often called "social" in a derogatory sense. They are not in principle limited to factors that are consensually considered as "theoretical" and "experimental" by practitioners at a given stage of the research.

To admit this is still not necessarily catastrophic, since one can maintain that the different constitutive ingredients of the symbioses have different weights. Pickering does not seem to be completely against this kind of explanation, given that he admits that the ingredients do not all have the same importance in the shift from one symbiotic state to another. However, he stresses that such an evaluation in terms of relative weights can be carried out only retrospectively, and that it has then a purely descriptive and local value: it cannot hope to be more than a simple *observation* that, in *this* particular case, certain factors have played a predominant role. It cannot be taken as an index of what had to be valid *de jure* in that particular case, and even less of what should in general be valid for an authentic science. Pickering endorses a strong contingentism, and from Pickering's contingentist point of view, nothing is absolutely predetermined. The actual weight of the factors in a given transition, once brought to light after the end of this transition, cannot be seen as already potentially contained in the past situation; it cannot be equated to a character contained in some initial conditions bound to lead to the final conditions that actually obtained (or to something reasonably close to them).

As for the question of the compelling power of the "quality of symbiosis" factor, and, correlatively, the question of its "degree of universality" (in the sense of the homogeneity with which it acts on the different practitioners), we can say that it depends heavily on the positions adopted with respect to the two preceding points. I confess that I find it difficult to identify Pickering's position with respect to the last problem. To talk about an "unforced and irreducible choice" clearly indicates that no absolute necessity is involved here. But who would claim, today, that such strong necessity is at stake in scientific investigation? Philosophers of science instead commonly invoke constraints, and under such retreat, the question becomes the force and legitimacy of the constraints involved. Now, Pickering's answer on the basis of the symbiotic model is not so clear.

9. CONCLUSION

Let us summarize what we have done.

Firstly, as a preliminary step, we have developed and justified our criteria for incommensurability verdicts: the kind of ACM involved must have damaging implications for at least one identifiable conception of science. The different situations discussed in the paper led to the conclusion that the associated issues are realism versus anti-realism and inevitabilism versus contingentism.

Secondly, we have analysed the idea of a machinic or literal ACM. We have shown that the condition that defines the machinic-literal incommensurability in Pickering's and Hacking's writings, namely "disjoint sets of machines or measures", is ambiguous. We have distinguished different ways to interpret it, depending on the conception we adopt of what an instrumental device and a measure is, and on the "criteria" according to which equivalent/non-equivalent classes of instruments and measures are constituted. We have stressed that in some cases, the definitions in terms of machines and in terms of measures are not at all equivalent: the fact of the matter one picks out could even be realized at the same times where the fact of the matter the other picks out is not.

Thirdly, we have shown that whatever may be the experimental disjunction between two experimental practices in detail, the very fact of an absence of intersection between the two corresponding sets of machines or measures is *not* sufficient for experimental incommensurability, that is, to load the factual experimental ACM with epistemologically harmful implications. We have identified two additional requirements. The first one is negative: it is the actual incapacity of practitioners to explain the two sets of experimental means and the measures obtained through them within one and the same theoretical framework. The second one is positive: the two experimental practices must be seen as competing practices, or at least as practices the ancestors of which have competed in the past. We have discussed the analogous requirement that holds in the case of traditional semantic incommensurability. From there we have made explicit a strong and problematic condition that should be satisfied by experimental practices in order to transform the experimental disjunction into an genuine experimental incommensurability: differences in contingent features of experimental practices, like the presence/absence of this or that instrumental device, should be seen as the cause of irreducible differences at the level of what is taken as a scientific fact in the broad sense, including assumptions about disciplinary mapping and kinds of reality.

Fourthly, we have shown that one of Pickering's central examples of experimental incommensurability is a special case of methodological incommensurability between local experimental standards. From this particular case we have given a general characterization of an experimental, methodological incommensurability. We have insisted that it would be useful, in studies devoted to experimental practices that flourished in recent decades, to recognize cases of the methodological kind as such, and to use the conceptual tools and the analyses that have been articulated in traditional discussions of the incommensurability problem.

Finally, we have provided sketchy and programmatic critical remarks about the machinic-literal incommensurability as a *local* characterization centred on some experimental items *only*. We have questioned the fitness of such a local characterization with Pickering's and Hacking's holistic ("symbiotic") model of scientific advance. We have suggested what might be a better characterization in the framework of the symbiotic model: a global incommensurability between scientific symbioses, namely, an incommensurability between two large packages of heterogeneous intertwined elements (including material ones) that altogether would constitute two self-consistent, self-referential and self-vindicating wholes.

The next steps would be to examine Pickering's other examples in the light of the previous developments, to analyse different possible kinds of experimental ACM in the transparent regime, and to articulate the symbiotic model of scientific advance.

BIBLIOGRAPHY

Ackermann, R. J. (1985) Data, *Instruments, and Theory: A Dialectical Approach to Understanding Science*. Princeton. NJ: Princeton University Press.

Bird, A. (2000) *Thomas Kuhn*. Princeton, NJ: Princeton University Press.

Bitbol, M. (2004) Néo-pragmatisme et incommensurabilité en physique. In L. Soler (ed.) *Le problème de l'incommensurabilité, un demi siècle après*. *Philosophia Scientiae,* 8(1), juin 2004, 203–234.

Carrier, M. (2001) Changing Laws and Shifting Concepts. On the Nature and Impact of Incommensurability. In P. Hoyningen and H. Sankey (eds.), pp. 65–90.

Feyerabend, P. (1975) Against *Method. Outline of an Anarchistic Theory of Knowledge*. London: New Left Books.

Galison, P. (1997) *Image and Logic: A Material Culture of Microphysics*. Chicago, IL: University of Chicago Press.

Hacking, I. (1992) The Self-Vindication of the Laboratory Sciences. In A. Pickering (ed.) *Science as Practice and Culture*. Chicago, IL: University of Chicago Press, pp. 29–64.

Hacking, I. (1999) *The Social Construction of What?* Cambridge, MA: Harvard University Press.

Hacking, I. (2000) How Inevitable Are the Results of Successful Science? *Philosophy of Science*, 67 (Proceedings), 58–71.

Hoyningen-Huene, P. and Sankey H. (eds.) (2001) *Incommensurability and Related Matters*. Dordrecht, The Netherlands: Kluwer.

Kuhn, T. (1991) The Road Since Structure, in PSA 1990. In A. Fine, M. Forbes and L. Wessel (eds.) *Proceedings of the 1990 Biennal Meeting of Philosophy of Science Association*, Vol. II. East Lansing, MI: Philosophy of Science Association, pp. 2–13.

Pickering, A. (1984) *Constructing Quarks: A Sociological History of Particle Physics*. Chicago, IL: University of Chicago Press.

Pickering, A. (1995) *The Mangle of Practice: Time, Agency and Science*. Chicago, IL: University of Chicago Press.

Sankey, H. (1994) *The Incommensurability Thesis*. Aldershot: Avebury.

Sankey, H. (1998) Taxonomic Incommensurability. *International Studies in Philosophy of Science*, 12, 7–16.

Soler, L. (2000) *Introduction à l'épistémologie*. Paris: Ellipses.

Soler, L. (2004) The Incommensurability Problem: Evolution, Approaches and Recent Issues, in Le problème de l'incommensurabilité, un demi siècle après/The Incommensurability Problem, Half a Century Later. *Philosophia Scientiae*, 8(1), 1–38.

Soler, L. (2006a) *Philosophie de la physique : dialogue à plusieurs voix autour de controverses contemporaines et classiques*, entre Michel Bitbol, Bernard d'Espagnat, Pascal Engel, Paul Gochet, Léna Soler et Hervé Zwirn, Léna Soler (ed.), Paris: L'Harmattan, collection 'Epistémologie et philosophie des sciences'.

Soler, L. (2006b) Contingence ou inévitabilité des résultats de notre science? *Philosophiques*, 33(2), 363–378.

Soler, L. (2006c) Une nouvelle forme d'incommensurabilité en philosophie des sciences? *Revue philosophique de Louvain*, Tome 104, n°3, août 2006, 554–580.

Soler, L. (2007), forthcoming. Revealing the Analytical Structure and some intrinsic major difficulties of the contingentist/inevitabilist issue. *Studies in History and Philosophy of Science*.

Soler, L. and Sankey, H. (eds.) (2007), forthcoming. Contingentism versus inevitabilism: are the results of successful science contingent or inevitable? A set of papers by Allan Franklin, Howard Sankey, Léna Soler and Emiliano Trizio, with an introduction by Léna Soler, *Studies in the History and Philosophy of Science*.

SOME REFLECTIONS ON EXPERIMENTAL INCOMMENSURABILITY

Commentary on "The Incommensurability of Experimental Practices: An Incommensurability of What*? An Incommensurability* of a Third Type*?", by Léna Soler*

HOWARD SANKEY

1. EXPERIMENTAL INCOMMENSURABILITY

Léna Soler discusses a purportedly new form of incommensurability, which she refers to as "machinic literal incommensurability". She also employs the expression "experimental incommensurability" to refer to roughly the same thing. As the former is a bit of a mouthful, I will opt for the latter expression.[1]

Experimental incommensurability relates to variation in equipment and associated experimental practices across different laboratory contexts. It arises because no common measure exists between different instruments and practices associated with these different instruments. Such an absence of common measure between different instruments and practices is supposed to constitute a new and distinct form of incommensurability. It contrasts with the standard semantic form of incommensurability which arises due to semantic variation between the vocabulary utilized by theories. It also contrasts with methodological forms of incommensurability which arise due to variation in the methodological norms employed for the evaluation of alternative scientific theories. Experimental incommensurability arises due to variation of equipment and experimental practice between experimental settings. It has nothing to do with the meaning or reference of scientific terms, nor has it anything to do with the norms of scientific theory appraisal.

Léna Soler argues that experimental incommensurability involves methodological incommensurability, and may even be a form of such incommensurability (see her discussion of Pickering in Sect. 7). As such, it does not constitute a new form of incommensurability. I think this is right. But I will attempt to broaden this result by showing that it is not a new form of incommensurability in addition to *either* semantic

[1] The difference between "machinic literal" and "experimental" forms of incommensurability is a subtle one. Soler says that the latter is more general (p. 300). Presumably, the former involves instrumentation, while the latter relates to experiment in general. The difference need not concern us here.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 341–347.

or methodological incommensurability. But, before arguing for this claim, I will deal with several other issues first.

2. INCOMMENSURABILITY AND EXPERIMENT

Before turning to the content of Soler's paper, I would like to make two brief remarks about the relationship between the idea of experimental incommensurability and the more traditional discussion of incommensurability.

It is of some interest to note that Kuhn made passing reference to the possibility of incommensurability relating to experimental apparatus in his original discussion in *The Structure of Scientific Revolutions*:

> Since new paradigms are born from old ones, they ordinarily incorporate much of the vocabulary and apparatus, both conceptual and manipulative, that the traditional paradigm had previously employed. But they seldom employ these borrowed elements in quite the traditional way. Within the new paradigm, old terms, concepts, and experiments fall into new relationships one with the other. (Kuhn, 1996, p. 149)

I would not wish to cite this passage as evidence that Kuhn anticipated the idea of experimental incommensurability. Nor would I wish to cite it as evidence that experimental incommensurability is not a new form of incommensurability, since it formed part of Kuhn's original notion of incommensurability. But I think it is evidence that, in his original way of thinking about the topic, Kuhn was sympathetic to the idea that difference in the use of experimental apparatus might be one basis for incommensurability.

Second, it is important to note that the idea of experimental incommensurability arises in the context of a specific development in recent history and philosophy of science. The traditional discussions of semantic and methodological incommensurability were situated in an earlier problematic. They took place in the context of the problem of scientific theory comparison, where philosophers were concerned to defend the rationality and progressiveness of scientific theory change and choice from various relativist and irrationalist challenges. By contrast, the idea of experimental incommensurability has arisen within the context of the "new experimentalism", which is proposed as a corrective to the traditional focus on scientific theories, theory appraisal and the relation between theory and reality. As a result, the location of experimental incommensurability tends to be scientific instrumentation and experimental practice, rather than evaluation or comparison of alternative scientific theories.

Thus, experimental incommensurability would arise in the first instance as a problem for the comparison of scientific instrumentation, experimental results and laboratory practice. But, since theories may be grounded in such instruments, results and practices, experimental incommensurability has the potential to give rise to an incommensurability between alternative theories as well.

3. CONDITION C: CATASTROPHIC CONNOTATIONS

Soler introduces two "minimal requirements" which must be satisfied for "legitimate use of the term 'incommensurability' " (p. 303). The first is the Absence of Common Measure condition. The second is Condition C: the catastrophic connotations condition. I regard the first condition as relatively unproblematic (though I will discuss a related issue in the Sect. 4). But Condition C is a retrograde step which is to be resisted.

According to Condition C, in order for use of the term "incommensurability" to be "legitimate", use of the term must give rise to what Soler describes as "catastrophic connotations". This is because the term "incommensurability" has "a strong emotional potential" (p. 303):

> [W]hat is the effect of the word "incommensurability" on the group of people who reflect on the nature of science …? Beyond individual fluctuations we can say, globally, that the term very largely induces strong, if not violent reactions. These reactions can be negative or positive. But whatever their quality may be, the term "incommensurability" is not neutral. (Soler, p. 304)

Soler explains that the reason why use of the term elicits strong reaction is that it implies that "something essential is questioned about science" (p. 303). Those whose reactions are negative find incommensurability to be "damaging" (p. 304) or "dangerous" (p. 305). Those whose reactions are positive find it to be "liberating" (p. 305). All parties agree that "something essential [about science] is at stake" (p. 305).

It is first to be noted that controversy has surrounded the idea of incommensurability almost since it came on the scene in the philosophy of science. Soler is therefore right that as a matter of the *reception* of the idea, incommensurability has been the subject of strong reactions. Nevertheless, the proposal to treat "catastrophic connotations" as a criterion for the application of the term is an ill-conceived proposal.

The "catastrophic connotations" of the idea of incommensurability are not the same as the idea itself. It is no part of the content of the claim of incommensurability that incommensurability has catastrophic consequences. This is so whether one is thinking in terms of semantic or methodological forms of incommensurability. To say that two theories or experimental practices are incommensurable is to say that they are incomparable either by means of shared semantic content or by means of shared methodological standards. While it may be controversial to say that theories or experimental practices are incomparable in these ways, it is no part of the claim of incommensurability that such incomparability must itself be considered to be controversial. That is something that depends on the way in which the claim of the incomparability of theories or practices is received by philosophers of science. Thus, the controversial character of the claim is something in addition to the claim itself, which depends on how the claim is received.

Soler appears to concede this point at the start of Sect. 3.3 where she allows that the "damaging consequences" of incommensurability "are not intrinsic features of the scientific configurations under study" (p. 305). But this is only an apparent concession. For, as we have seen above, Soler treats Condition C as a criterion for the application of the term "incommensurability", since she says that it is a "requirement" for the "legitimate use of the term" (p. 303). This means that if a particular application of the term fails to have "catastrophic connotations", it fails in that instance to denote a genuine case of incommensurability. That is, because Condition C is a criterion for the application of the term "incommensurability", Condition C erroneously makes the "catastrophic connotations" part of the very content of the idea of incommensurability.

Failure to distinguish response to the claim of incommensurability from the content of that claim does little to improve the clarity of the discussion. But, as I will now argue, the situation is worse than this. It leads to the unwelcome result that paradigm cases of claims of incommensurability may fail to count as instances of the incommensurability thesis.

I will illustrate the point with reference to Kuhn's later taxonomic version of the incommensurability thesis. In his later work, Kuhn proposed a refined version of the semantic incommensurability thesis, according to which incommensurability consists of failure of intertranslatability of localized subsets of interdefined terms within the lexicons of alternative theories (see Kuhn, 2000, Chap. 2). Because such local incommensurability is restricted to subsets of terms, there is significant scope for the direct comparison of the empirical predictions of theories because these may be expressed in vocabulary that is shared between the theories. Kuhn is careful to distinguish between translation and communication, insisting that local incommensurability presents no obstacles for communication or understanding between scientists. And, as I have shown in my paper, "Taxonomic Incommensurability" (Sankey, 1998), the localized incommensurability thesis represents no threat whatsoever to a scientific realist account of scientific progress.

In short, Kuhn's final taxonomic incommensurability thesis represents no threat whatsoever to the rationality or progressiveness of science. Its "emotional potential" of both a negative and positive character has therefore been downgraded. Its "catastrophic connotations" have been removed. But, so far from saying that such taxonomic incommensurability fails to qualify as a species of incommensurability, as Soler must do, what follows is that the taxonomic incommensurability thesis is a weakened version of the incommensurability thesis from which much of the interest has disappeared.

4. THE PROBLEM OF RIVALRY

As noted at the start of the Sect. 3, I regard Soler's Absence of Common Measure Condition as relatively unproblematic. The condition is well-motivated given the meaning of the word "incommensurability". For there to be no common measure is part of what it means to be incommensurable.

But problems emerge when one probes further into the relationship between incommensurable alternatives. In Sect. 6.2.2, Soler claims that "we must include the *requirement of rivalry* as a *necessary condition* of the experimental incommensurability verdicts" (p. 318). Soler further expands the point as follows:

> When we say that two scientific practices compete, we assume, at the most fundamental level, that they are mutually exclusive in principle and hence that both should not be maintained in fact. This goes with the demand to choose either one or the other, or alternatively, to find a way to reconcile them.… If we … ask what kind of assumption lies behind the idea that two scientific practices are mutually exclusive and cannot both be maintained, we find the commitment that the two researches under comparison should not be in conflict *because they study the same reality*: for example the physical world for two competing physics.… Let us speak of a rivalry related to the identity of the targeted object. (p. 318)

In this passage, and throughout much of Sect. 6, Soler's aim is to show that a condition of rivalry must also be imposed on claims of experimental incommensurability. In order for the claim of incommensurability to get a grip, it must be the case that the purportedly incommensurable practices are in fact in a relation of rivalry.

The requirement of rivalry has a sound rationale that has been widely endorsed in the semantic incommensurability literature. It reflects the fact that the problem of incommensurability arises in relation to choice between competing alternatives. The usual context for discussion of both semantic and methodological forms of incommensurability has been that of scientific theory-choice, in which scientists must choose between two or more competing theories, for example, in the context of a scientific revolution. In the case of experimental incommensurability, the context is presumably one in which a choice is to be made between alternative experimental practices or scientific instrumentation.

But to treat rivalry simply as a requirement obscures a significant fact about the issue of rivalry. The notion of rivalry originally entered discussion of the topic of incommensurability as the basis for an objection to the idea of incommensurability, rather than as a requirement on the relation of incommensurability. Authors such as Dudley Shapere (1966) and Israel Scheffler (1967, especially p. 82) objected that it was entirely unclear how theories with no meaning in common, without common observations or shared evaluative standards, and which might even be about different Kuhnian "worlds", could ever enter into a relation of rivalry in the first place. The problem was particularly acute in relation to the issue of semantic incommensurability. For if terms of incommensurable theories share no meaning in common, then it is not possible for assertions of one theory to either assert or deny assertions of a theory with which it is incommensurable (Shapere, 1966, p. 57; Feyerabend, 1981, p. 115).

The proposal to treat rivalry as a requirement for incommensurability fails to address the problem of how a relation of rivalry can even arise between incommensurable alternatives. In the passage cited above, Soler does mention that incommensurable alternatives are "mutually exclusive" and notes that there must be an "identity of the targeted object". But, in the context of the issue of semantic incommensurability, it remains unclear how assertions stated in semantically variant terminology might be "mutually exclusive", since the absence of shared meaning entails that they are unable to enter into a relation of conflict. And being about the same "targeted object" does not entail rivalry, since assertions about the same thing do not necessarily enter into disagreement.

These problems have never been entirely resolved by proponents of the thesis of incommensurability. And it is not without interest to remind ourselves of these problems in the context of the purportedly new experimental form of incommensurability. For how exactly are different sets of instruments, or different experimental practices, meant to enter into a relation of rivalry with one another? To merely insist upon rivalry as a requirement for experimental incommensurability will not take us very far. For what remains to be explained is how different sets of instruments or practices might enter into a relation of rivalry in the first place.

What I suggest is that the only handle we have on the relevant notion of rivalry involves the notion of disagreement. The condition of rivalry will obtain when assertions about instruments or practices enter into conflict. This is the only form of rivalry of which we have any clear understanding. Given this, I suggest that the problem of experimental incommensurability reduces to the problem of semantic incommensurability. Until such time as we have reason to believe in semantic incommensurability, we have no reason to take talk of experimental incommensurability seriously. I will attempt to reinforce this claim in Sect. 5.

5. A NEW FORM OF INCOMMENSURABILITY?

I turn now to the question of whether experimental incommensurability is a form of incommensurability distinct from either semantic or methodological forms of incommensurability. For the sake of argument, let us assume that the idea of experimental incommensurability is sufficiently clear that it makes sense to assume that there may be distinct experimental practices grounded in different instrumentation. Given the differences between the instruments employed, there might be no way to directly compare the results derived from experiments using these instruments. I take it that this is the sort of situation in which experimental incommensurability is supposed to arise.

Now, if experimental incommensurability is a distinct form of incommensurability, then it ought to arise even in the absence of semantic and methodological forms of incommensurability. That is, it should be possible for alternative theories, practices or instruments to be machine incommensurable even if they are semantically or methodologically commensurable.

But if this is so, then it is not clear that any real problem of commensurability arises. If vocabulary and concepts are shared, then there is no problem with comparison of content, since theories, practices and experimental results can be compared using a common language. And if the norms of scientific research are also shared, there can be no methodological obstacle to the comparative appraisal of theories, practices or equipment.

Let me try to make this point clear by means of a simple example. Consider a simple instrument, the ruler, or measuring-stick. Rulers are used to measure length. Different systems of measurement are used in different parts of the world. Some of us employ the imperial system of feet and inches, while others use the metric system. Rulers can be based on either system of measurement. There is therefore a sense in which those who use an imperial ruler use a different instrument of measurement from those who use a metric ruler.

Since those who use the imperial ruler to measure length will arrive at a different numerical result from those who use a metric ruler, the question arises of whether any common measure exists between the two instruments. In a quite literal sense, the question arises of whether the two measuring instruments are commensurable.

One might seek to establish a common measure between the two instruments by employing each of the rulers to measure the units of measurement on the other ruler. If we proceed in this way, we would find, for example, that an inch equals 25.4 mm, and that a millimetre equals 0.039 of an inch. We could then set up a conversion from one system into the other, so that measurements produced using one ruler could be converted to measurements that would be obtained using the other kind of ruler. If such conversion is possible, then it is possible to compare measurements obtained using one kind of ruler with measurements obtained using the other kind of ruler.

But suppose it is impossible to convert from one system of measurement into the other. Then it might be said that the imperial and metric rulers were incommensurable. But such incommensurability would be incommensurability in the standard, semantic sense of inability to compare due to failure of translation.

The only other possible source of incommensurability would be if there were no common procedures of measurement. Suppose that users of the imperial ruler only allow direct visual inspection of the ruler, whereas the users of the metric ruler only permit tactile inspection of the ruler. Then there would be no common procedure of measurement. But if there is no common procedure of measurement, then the incommensurability that arises would be methodological incommensurability due to lack of common methodological norms or standards.

At least in the case of this simple example, there is no apparent basis for postulating the existence of a third, experimental form of incommensurability, over and above the standard semantic and methodological forms of incommensurability. The ball therefore lies in the court of those who think that there is a further, distinct form of incommensurability that needs to be introduced.

BIBLIOGRAPHY

Feyerabend, P. (1981) Reply to Criticism: Comments on Smart, Sellars and Putnam. In *Realism, Rationalism and Scientific Method: Philosophical Papers, Vol. 1*. Cambridge: Cambridge University Press, pp. 104–131.

Kuhn, T. S. (1996) *The Structure of Scientific Revolutions*, 3rd ed. Chicago, IL: University of Chicago Press.

Kuhn, T. S. (2000) *The Road Since Structure*. In J. Conant and J. Haugeland (eds.). Chicago, IL: University of Chicago Press.

Sankey, H. (1998) Taxonomic Incommensurability. *International Studies in the Philosophy of Science,* 12, 7–16.

Scheffler, I. (1967) *Science and Subjectivity*. Indianapolis, IN: Bobbs-Merrill.

Shapere, D. (1966) Meaning and Scientific Change. In R. G. Colodny (ed.) *Mind and Cosmos*. Pittsburgh, PA: University of Pittsburgh Press, pp. 41–85.

PART 12

PRAGMATIC BREAKDOWNS: A NEW KIND OF SCIENTIFIC REVOLUTION?

DISRUPTIVE SCIENTIFIC CHANGE

THOMAS NICKLES[1]

Abstract Much of the philosophical treatment of incommensurability is an artifactual response to internally generated philosophical problems rather than to the difficulties faced by scientists themselves. As a result, incommensurability and rupture have been mislocated at the level of scientists' differing beliefs and their disagreements about symbolic representations of nature, and that of corresponding failures to meet the foundational demands of one or another philosophy of language. Kuhn famously attacked the view that scientific progress is cumulative, but his revolutions – representational ruptures – are rare events, the sometime products of extraordinary science that threaten to bring inquiry to a halt. Whereas, when we examine scientific practice at the frontier, even in "normal science," we find that scientists not only cope but also thrive on the unstable ground found there. Theory rupture is causally neither necessary nor sufficient for the disruption of scientific practice. Indeed, alternative practices are sometimes the stimulus for radical theory change. We can read Kuhn himself as emphasizing practice over representations and (certainly) rules, but Kuhn's discussion is limited to a few cases, pre-World War II. In the final sections I invite consideration of alternative patterns of disruption by comparing some scientific examples with non-Kuhnian sorts of disruption that occur in business life – which, like science, is highly competitive and places a premium on innovation.

Keywords disruption, heuristic appraisal, incommensurability, Kuhn, professional bias, scientific discovery, scientific practice, scientific revolution.

> Why do so few "scientists" even look at the evidence for telepathy, so called? Because they think, as a leading biologist, now dead, once said to me, that even if such a thing were true, scientists ought to band together to keep it suppressed and concealed. It would undo the uniformity of Nature and all sorts of other things without which scientists cannot carry on their pursuits. But if this very man had been shown something which as a scientist he might *do* with telepathy, he might not only have examined the evidence, but even have found it good enough.
>
> —William James (1896, Sect. iii)

[1] First, my thanks to Léna Soler, Howard Sankey, Paul Hoyningen, the referees, and the sponsors of the conference, which took place in the beautiful city of Nancy. I even found a little street named Nickles near the university. Thanks also to comments from the other participants and especially to Emiliano Trizio for his helpful commentary. In Reno, Yoichi Ishida provided good advice.

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 351–379.

> To these strokes of boldness [of the new relativistic mechanics of Lorentz], however, we wish to add more, and much more disconcerting ones. We now wonder not only whether the differential equations of dynamics must be modified, but whether the laws of motion can still be expressed by means of differential equations. And therein would lie the most profound revolution that natural philosophy has experienced since Newton … [the rejection of the principle] that the state of a system in motion, or more generally that of the universe, could depend only on its immediately preceding state; that all variations in nature must take place in a continuous manner.
>
> —Henri Poincaré (1913)[2]

1. INTRODUCTION

Most treatments of scientific revolutions and incommensurability are either theory-centered or methodological (Hoyningen-Huene and Sankey, 2001) – or both. Accounts of the first kind take the representational content of science to be primary and focus on displacements in the logic, meaning, or reference of theories during major theory change. Those of the second kind draw our attention to incompatible goals, differing value commitments, and the correspondingly diverse modes of evaluation. Both approaches typically have an epistemological emphasis. Good theories express the truth, or are closer to the truth than their predecessors; and the important methodological requirements are truth-conducive.

My approach will be different. Rather than representational and epistemic, it will be pragmatic and heuristic, i.e., practice-centered and future-oriented. I shall argue that there is a double asymmetry in the importance of incommensurability as it concerns philosophers and scientists. On the one hand, many philosophers and some intellectual historians regard *conceptual* breaks (major representational changes) as earth-shaking, as if such changes threaten to stop scientific research in its tracks. By contrast, scientists themselves frequently adapt to them easily and often find them highly stimulating. The more creative scientists actively seek them. On the other hand, major alterations in scientific *practice* can and do disruptively transform scientific communities, yet most philosophers pay little attention to changes of this type.

One reason for these asymmetries (I claim) is that philosophers and scientists have different professional commitments or biases (in the non-pejorative sense of "bias"). Such differences are natural and legitimate, up to a point; but, given that one task of philosophy of science is to understand how the sciences themselves work, it is distressing how many philosophers' problems have been artifactual, generated by their own inadequate philosophies of language or by priggish commitments to conservative philosophical principles (Shapere, 1984, 2001) – by contrast with the unprincipled, pragmatic opportunism of the better scientists. As Einstein famously remarked, "The scientist … must appear to the systematic epistemologist as an unscrupulous opportunist."[3] Matters become still more complicated if we take into account also the professional biases of historians and sociologists of science.

2 Poincaré (1963, p. 75), "The Quantum Theory." Apropos of the James quote (from "The Will to Believe"), see Ruse (2001).

3 In Einstein's reply to Lenzen and Northrop, in Schilpp (1949, p. 684).

Consider intelligibility. The logical empiricists understood this in terms of a theory's logical consistency, conceptual coherence, and cognitive significance. Inconsistency was the worst sin imaginable for them. We can understand why, given that the first-order predicate calculus, with its paradox of the material conditional, was their canonical language for logical reconstruction; for from an inconsistency any claim at all immediately follows. But this logical blowup was an artifact of their formal methodology, a result that makes it puzzling how scientists can usefully employ models that are inconsistent or incoherent, as they often do (e.g., Rohrlich, 1965; Shapere, 1984, Chap. 17; Frisch, 2005; Arabatzis, 2006). As for truth, an overriding concern with true belief and sound arguments as the product of scientific work makes puzzling the pervasive use of models in the sciences, given that models, almost by definition, are deliberately flawed in certain respects. At least as puzzling is how scientists can use deliberately oversimplified models in order to reveal more fundamental truths. For what could be the point of multiplying mutually contrary models known from the start to be false? Yet as biologist Richard Levins famously said, "our truth is the intersection of independent lies" (Levins, 1966, p. 20; Wimsatt, 1987).

Now there *is* something that could be called unintelligibility within scientific work besides an inability to make sense of a claim or a practice. It is an inability of scientific specialists to see the *relevance* of a research claim, technique, or strategy to their field, to their expertise, to their own specific work. As scientists, they cannot figure out anything important that they can *do* with it. It does not engage their practices. It is not a tool that they can use to formulate or solve problems, or at least not one that they consider reliable and legitimate to use. A solid state physicist may say of such a claim or practice: "That's not solid state physics!" A proteomist may say: "That isn't proteomics!" In short, these specialists fail to *recognize* it as a contribution to their field. This kind of intelligibility problem is a recognition problem or identification problem.[4]

Something is *unintelligible-as-X* in my sense if it does not mesh with the X community's expert practices, their know-how. Intelligibility, transparency, depends upon practical expertise. Failure to agree at the levels of theoretical articulation and methodological rules does not necessarily interrupt agreement over problem-solving practices, as Thomas Kuhn himself emphasized (Kuhn, 1970, Sect. V and Postscript). Such disagreement does not signal rupture. It is the problem-solving practices that count most.

One sort of rupture is occasioned by widespread failure to master the practices in question. Another is failure to see the point or the promise of those practices. In the late nineteenth century, an old-style, experimental thermodynamicist might be unable to apply the new-fangled statistical-mechanical techniques, but he might also simply regard the new approach as pointless, sterile, unpromising. In this extended sense, it is unintelligible as a *positive contribution* to his field as he practices it. Of course, a serious rupture must affect most of the practitioners in a field. We can speak of individual recognition failures in the context of personal biographies, but we are here concerned with communities of scientists.

[4] As my astute commentator Emiliano Trizio pointed out, there are kinds and degrees of practical (un)intelligibility that need to be distinguished (more carefully than I can here) See also Nickles (2003a).

When the topics of scientific revolution and incommensurability come up, the discussion naturally turns to Kuhn and Paul Feyerabend. I shall focus on Kuhn. Perhaps his greatest merit was to point out that there are different tempos and modes of scientific work, and that scientific advances need not be cumulative. Speaking of revolution was his way to emphasize the sometimes disruptive character of scientific work. Yet with today's hindsight it seems fair to say that only against the traditional philosophical background and the strong, scientistic whiggism of writing about science could Kuhn's claims be so shocking. For what should we expect at the frontier of research? It is here that the problem of induction, in its forward-looking form, has real bite. For even the experts do not know whether, or how well, previously gained results and techniques can be projected onto the new domain. And from this perspective, we suddenly realize that Kuhn did not go far enough, for *his* episodes of non-cumulative progress occur less frequently than scientific disruptions in general.

However, in my opinion there are kinds of disruptive change that do not fit Kuhn's account very well. His treatment of revolutions fails to capture the nineteenth-century transformation of physics into modern, mathematical physics, for example (see Sect. 6). And it is arguable that science (many sciences) have entered a new cultural phase since World War I, especially since World War II, a phase that philosophers, in contrast to historians and sociologists, have been slow to appreciate.[5] During this period we find specialty areas popping into and out of existence, recombining and transforming themselves in all kinds of ways; and we have "little science" becoming "big science." It is as if an account of invention and technology stopped with the era of individual inventor, neglecting the era of big industrial, biotechnological, and university laboratories.

Essential novelty in itself creates recognition problems. According to the (possibly apocryphal) story, the young physicist gave a lecture concerning a current problem at Bohr's institute in Copenhagen. Taking a walk with him afterwards, as was Bohr's custom, Bohr remarked, "I think your ideas are crazy. The trouble is, I do not think they are crazy enough!" We understand the point of the story. There are times when radical new ideas or practices seem necessary to break through an impasse, namely, those occasions in which all conventional approaches seem hopeless. But we also appreciate the danger. In the arts as well as the sciences, insofar as something is iconoclastic, other practitioners may not recognize it as a contribution to the field in question. Whoever heard of creating a painting by dripping and splashing paint on a horizontal canvas, laid flat on the ground?

[5] William James caught a whiff of the coming changes more than a century ago when he wrote (of men of genius):

> Instead of thoughts of concrete things patiently following one another in a beaten track of habitual suggestion, we have the most abrupt cross-cuts and transitions from one idea to another, the most rarefied abstractions and discriminations, the most unheard of combination of elements, the subtlest associations of analogy; in a word, we seem suddenly introduced into a seething cauldron of ideas, where everything is fizzling and bobbling about in a state of bewildering activity, where partnerships can be joined or loosened in an instant, treadmill routine is unknown, and the unexpected seems only law. (James, 1880, p. 456, 1956, p. 248)

Kuhn (1977, pp. 166 ff.) made this point about normal science. Essential novelty is not wanted, he said; and when it does occur in the form of a confirmed discovery, it is a surprise – and somewhat threatening. If it is too radical, it will be dismissed as nonsense or simply ignored. This is the key to understanding his "essential tension" between innovation and tradition. Science places a premium on new results – but not so new that they fall completely outside extant research traditions. Essential novelty is either recognized or not. If recognized, it is disruptive; if not, it is simply disregarded (There are some striking historical cases!) – raising the question for historians of why scientists failed to appreciate it.[6]

Since Kuhn himself writes as a philosopher and an internalist historian as well as a trained physicist, tensions of another sort occur within his own work. He has a foot on each side of the double asymmetry. The reader must be constantly alert as to whether Kuhn's voice at a given point is that of Kuhn the historian, the philosopher, or the scientist. His discussion of incommensurability is limited, I believe, in that his serious scientific examples end with relativity and the old quantum theory.

Kuhn's inattention to the sciences after the World Wars leads him to construe incommensurability as something that happens within a single field and that centrally involves a theory (but see Sect. 4 for an alternative reading). There are only two kinds of Kuhnian revolution, the one that initiates the field as a mature science and the later, major transformations of that field such as the relativity revolution in mechanics. There is little of the colonization and cross-fertilization characteristic of recent science and, arguably, much previous science as well.

Where can we look for alternative models of deep or disruptive change? We do, of course, study history of science and science studies accounts of it; but it may stimulate the imagination to look at other fields as well. Toward the end of this chapter, I briefly present some economic models of deep change that differ from Kuhn's. Now I am one of the last people to want to apply capitalistic business models to everything, but it is helpful to examine other areas that are also highly competitive, that place a premium on innovation, that are responsive to new technology, and that have apparently entered a new phase since the world wars. The Austrian–American economist Joseph Schumpeter (1942) characterized a capitalistic economy as a process of "creative destruction." To what extent does this label apply to scientific work?

My next two sections sketch a defense of my asymmetry claim. Sect. 2 describes differences in professional biases, and Sect. 3 itemizes ways in which creative working scientists cope with the kinds of breaks that bother philosophers. Sect. 4 is devoted to two interpretations of Kuhnian revolutions, representational and practice-based. Sect. 5 presents the just-mentioned alternative models of disruptive technology. Sect. 6 considers some possible scientific analogues, and Sect. 7 returns to Kuhn's theme of the invisibility of revolutionary disruptions and presents a concluding irony.

[6] This is the debate over whether there is such a thing as a postmaturity in scientific discovery, as proposed by Harriet Zuckerman and Joshua Lederberg (1986). See the interesting discussions of prematurity and postmaturity in Hook (2001).

2. PROFESSIONAL BIASES

Philosophers (and the historians of science who have influenced them) possess professional biases different from those of scientists. That is not surprising, for all investigators have their professional commitments. However, given the differences in biases, the philosophers' bias can get in the way when the topic becomes science itself and how it works. Further trouble stems from the fact that present-day philosophers of science, despite rejecting logical empiricism, still have not fully escaped from the conditions of origin of the field. Brevity forces me to resort to generalizations that require more qualification than I can furnish here, given that philosophers, historians, and scientists come in many shapes and sizes.

Creative scientists, as the Einstein quote suggests, are often pragmatic opportunists. There are zillions of historical examples of scientific theories, models, etc., that scientists have found fruitful to develop despite their initial logical and semantic flaws. Their heuristic appraisal (expected fertility) was high even though their epistemic appraisal (likelihood of being approximately true) was rock-bottom (Nickles, 2006b). This lack of internal coherence is characteristic not only of nascent theorizing but also of many exemplary scientific publications, including all three of Einstein's most famous papers of 1905, on the photoelectric effect and free quanta, Brownian motion, and special relativity. Given their emphasis on logical coherence, it is surprising how little the logical empiricists made of this fact.

What I want to emphasize here is that scientists are quite accustomed to working under these conditions and even seek them out. Nor should this surprise us. When we stop to think about it, nearly all hypotheses, theories, and models of any depth are initially ill-formed in various ways. By their nature they venture into the unknown, going beyond current resources in some way — although they must be constructed from those very resources.

One reason why philosophers still overlook this elementary fact, I believe, is that the old prohibition about discussing context of discovery lingers. In addition, the neglect of heuristic appraisal – the evaluation of future prospects of alternatives already on the table – reinforces this blindness to scientists' intellectual agility. Underlying both of these is the failure to take scientific practice seriously. Recently, philosophers have finally attended to experimental practices but not as much to theoretical practices. Philosophers still tend to start from scientific work that is "finished" and in the textbooks or classic papers.

What, then, are the professional biases of creative scientists? The primary one is wanting to get on with fruitful work so as to make the greatest contribution of which one is capable. A scientist is a forward-looking problem solver, active at the frontier of research. It is not enough to be a science teacher, treating science as a body of established knowledge to be conveyed in the form of a catechism. A young postdoc wants to find problems that she can solve and publish within the post doc years – the most challenging research that she is capable of doing, and to have her next project underway while finishing the current project; better yet, to have several projects going at once. After all, if she stops *doing* science, then, tautologically, she is no longer a working scientist! (As Descartes might have put it, qua scientist: "I research; therefore, I am.") Until their research careers wane, productive scientists rarely have time to stop and reflect systematically, in a philosophic and aesthetic mode, on the big frameworks their discipline may suggest.

A creative scientist who regards a theoretical model as an active research site treats it as the theoretical analogue of an experimental system – not as an aesthetic object to be admired as filling in part of our big picture of the universe but as a research tool and as itself an *object* of research, a system to be tinkered with in order to see what interesting results it might yield. ("What happens when we do *that* to the theoretical system?") Like experimental systems, theoretical systems can be full of uncertainties and surprises, unexpected hints but also resistances (Rheinberger, 1997; Arabatzis, 2006). And once they are developed to the point of fairly routine work (topics for graduate students to pursue), creative scientists move on to greener pastures, perhaps now using the polished model as an established research tool.

Mine is not the cynical position that truth is unimportant to scientists, nor that they are unconcerned about world pictures. My point is that, as far as their research goes, truth and pictures are more important for their heuristic value than for anything else. Truth claims about representations of the universe serve as pointers away from blind alleys and toward fruitful problems for future research (Nickles, 2006b). They also function precisely to help scientists recognize or identify other frontier work, divergent though it might be, as contributions to the "same" topic.

Philosophers often miss these functions. Philosophers' confirmation theories are backward-looking, purely epistemic assessments of the "track record" of a theory in order to decide whether or not to accept the theory as true, or approximately true, rather than forward-looking assessments of its heuristic fertility as a basis for further research. They are the philosophy-of-science version of the epistemologists' analyses of "The cat is on the mat." Meanwhile, confirmation in science is usually not absolute but comparative, attempting to show that one hypothesis or model fits the data better than another, if only the null hypothesis. Confirmation theorists have noted this but usually within discussions of truthlikeness. And confirmation theorists rarely address the important issue of deciding which among alternative research practices to adopt.[7]

I shall have more to say about scientists in Sect. 3, but let us now, again speaking quite generally, note how the professional biases of philosophers contrast with those of scientists. Philosophers tend to focus on the theoretical products of scientific research rather than on the process that produced those products. Philosophers want to know Reality before they die! Thus philosophical realists want to think that current theories are getting close to the truth. They aim at world views, not merely pragmatic problem solutions (and, in that respect, they are somewhat closer to religious advocates than to opportunistic scientists). They look to science as a better-grounded replacement for metaphysics. Since metaphysics aimed primarily to provide a comprehensive *representation* of the cosmos at its most fundamental, many philosophers apparently tacitly assume that that is also the primary purpose of scientific work. In the words of the classic pragmatists, Peirce, James, and Dewey, the stance of philosophers toward scientific theories and models is typically one of "spectators" who treat theories

7 For an extended case study, see Galison (1997).

as representations and sometimes as objects for aesthetic admiration rather than as research tools. Fortunately, this is now changing in specialty areas such as philosophy of biology and philosophy of cognitive science. Notice, by the way, that it is the kind of work done by philosophers of biology and cognitive science that the corresponding scientists find most useful – because the philosophers in these areas contribute to frontier science rather than focusing most of their attention on what scientists themselves consider meta-issues, that is, "merely philosophical" issues.

We philosophers of science remain captives of our professional history. For my account of why the logical empiricists and some of their immediate successors found formal logical and linguistic analysis so congenial, the reader may consult Nickles (2002). Basically, their task was to carve out a new field, one that possessed academic integrity but one autonomous from the empirical sciences themselves and from historical and sociological studies. Methodology of science conceived as "the logic of science" fit the bill admirably. Philosophy of science was to be a priori and analytic, not empirical. Rudolf Carnap emerged as the leader of this approach, and he and his followers marginalized Otto Neurath's very different approach, especially after the latter's early death. Since there appeared to be no logic of discovery, these philosophers attempted to fashion the tools that they believed were needed for logic of justification – to polish up the logical structure of scientific results already on the table. To adapt a famous saying of Kierkegaard ("We live forward, but we understand backward."), the logical empiricists were more concerned to "understand backward" than to investigate how scientists "live forward." Looking at actual scientific practice in detail was not on the philosophical agenda for two reasons. For one thing, that sort of empirical work was what historians, sociologists, and psychologists did. For another, there seemed to be no need, since the logical empiricists themselves had been trained in science and intuitively knew how research was conducted. Most of their work had been in physics and mathematics, so it is not surprising that, under the banner of "the unity of science," physics became the normative model for all sciences.

There were (and are) longer-term influences at work on philosophy of science as well. In my opinion the influence of Descartes and Kant remains a problem. When it comes to methodology, philosophers have often been priggish moralists, seeking epistemic foundations and laying down normative rules for everything, by contrast with the agile opportunism and risk-taking of scientists. Here I agree with Gaston Bachelard (1934) that general methodology as espoused by philosophers is nearly always conservative and, if taken seriously, a potential obstacle to future scientific development.

During the 1960s and 1970s, the new historians of science began to question the applicability of the methodological norms of the philosophers to real science as practiced. Their most serious criticism was that the norms were so obviously irrelevant to real science that enforcing them would kill science as we know it. This was a convincing answer to the reply that norms cannot be refuted by empirical facts, for if "ought" implies "can," then "cannot" implies "not ought." Many analysts concluded that historical case studies had refuted the normative methodologies of the logical empiricists and Popperians. Subsequently, some of the new sociologists of knowledge questioned the integrity and autonomy of philosophy of science altogether and attempted to close up the intellectual niche that the philosophers thought they had opened. A somewhat similar conclusion could be drawn from the return of naturalism within philosophy. For the discovery-justification and analytic-synthetic distinctions, together with the criterion of cognitive significance, had in effect demarcated philosophy, as a purely analytic enterprise, from the surrounding empirical disciplines. And according to the naturalists, there can be no purely analytical enterprise!

It was the new history of science, in the hands of N. R. Hanson, Stephen Toulmin, Kuhn, Feyerabend, Imre Lakatos, Larry Laudan, Dudley Shapere, and others, that created the battle of the big methodological systems. Those were exciting times indeed, and talk of scientific revolution and of deep conceptual change was at the center of it. We philosophers were carried away by our enthusiasm and missed the fact that working scientists were not bothered by the same kinds of things that we were. In other words, it was not just the positivists: historical philosophers of science also possess professional biases, professional "parameter settings" that differ from those of scientists!

What, then, of the professional biases of historians? The internalist, intellectual historians were responsible for "hyping" revolution! Historians and political scientists had long discussed social and political revolutions, but there had been scarcely any talk of scientific revolution until history of science began to emerge as a professional specialty area. Today's readers are often surprised to learn that hardly anyone spoke of epistemological ruptures, scientific revolutions, or even "the Scientific Revolution" until Gaston Bachelard (1934) and Alexandre Koyré (1939) began to do so in the 1930s. Herbert Butterfield's little book, *The Origins of Modern Science* (1949), popularized the idea of scientific revolutions, a theme that Kuhn amplified in *The Copernican Revolution* (1957) and *The Structure of Scientific Revolutions* (1962, 1970).

Why, then, the recent attraction of revolutions for historians? Here is my suggestion. Keep in mind that although the three social forces that have done more than anything to shape modern society are the development of capitalistic business life, the evolution of parliamentary democracy, and the emergence of modern science and technology, most people find these topics and their history dismally boring. After all, aren't scientists just fact-grinding drudges in white coats?

Thus the emerging discipline of professional history of science faced its own issues of integrity, autonomy, and relevance.[8] Unlike most twentieth-century philosophers, historians have sought an audience beyond that of other professional historians, indeed beyond academe, but how could an account of technically esoteric scientific work be made to appeal to a wider public and to address the big questions of the day? Enter the idea of scientific revolution!

The revolution theme immediately suggests many attractive narrative forms as well as helping historians organize the vast pile of historical information at their disposal. Revolutions have a beginning, a middle, and an end. They are sufficiently episodic to be dramatic, emotion-arousing. Revolutions explode the myth that scientists are drudges. They invite the creation of heroic leaders as well as villains, where the

[8] Nickles (2006a) provides a detailed list of ways in which past historians have exploited the revolution theme.

leaders were once obscure underdogs challenging orthodoxy. Biography therefore seemed an appropriate genre, where the job of the biographer (including graduate students writing dissertations) was to employ contrastive techniques to make even minor work seem as revolutionarily important as possible. Moreover, individual biographies are always serendipitous, so it was natural to see science in the large as more serendipitous than methodical. This in turn suggested that the history of science, as we know it, was not inevitable but, to some degree, a series of contingencies, historical accidents frozen in time. On such a view, historical narrative rather than general methodology is clearly the preferred way to understand science. Once again we have the ideographic pitted against the nomothetic.

Although the internalist, biographical approach soon gave way to a more social-constructivist approach, the themes of contingency and the freedom or agency of scientists remained central. On this view scientific work is not immune from the contingencies and choices that pervade everyday life. It is not rigidly constrained by a "scientific method." Stated in terms of professional biases, the point is that, whereas philosophers of science have (with some exceptions) tried to make scientific work culturally special and epistemologically privileged, in virtue of its special goals and methods, historians and sociologists of science have tried to construe scientific work as ordinary, characterized by contingent situations and choices.

Every field has its professional commitments, but these can be distorting when one field claims to regulate, or even merely to interpret, another. The opportunism point is also profession-relative. Philosophers, too, are opportunistic within their profession, but that does not encourage them to recognize the professional opportunism of scientists. Some degree of pragmatic opportunism is surely characteristic of the more creative members of all professions. It is sociologists of science who have most explicitly recognized this as well as the reflexive boomerang upon themselves.

3. HOW DO SCIENTISTS COPE WITH (REPRESENTATIONAL) CHANGE?

The deep metaphysical change already apparent in the early quantum theory that Poincaré noted in one of his last essays (in the quotation at the beginning of this paper) was indeed disconcerting to scientists and philosophers, who thought they already possessed a good grasp on reality. While I don't deny the seriousness of the representational break, I suggest that the unsettlement was not purely representational and epistemic. Rather, it was because the change in view disrupted scientists' deepest intuitions about how to proceed in physical research, intuitions backed by more than two centuries of success in applying mathematical analysis ("the calculus") to scientific problems.

We can distinguish two kinds of skepticism about epistemological realism here. Culture-carrying scientists (as well as philosophers) concerned about establishing a definite conception of reality must worry about the metaphysical disruptions that can be attributed to deep scientific change (Laudan, 1981), but other scientists are pragmatic problem-solvers who are not very concerned about metaphysical reality in the first place. Such scientists, then and now, were/are more concerned with whether or not the familiar differential equations would have to be replaced and, if so, with what and how fruitfully. Surely they did worry about whether the worldview of contemporary physics was wrong, but they worried even more (I claim) about the possibility that, procedurally speaking, classical physics had been on the wrong track. They worried about how much of their previous expertise would remain useful under the new dispensation. To many scientists, the epistemic status of a claim is as important for its heuristic value as for anything else. Elsewhere (Nickles, 2006b) I argue that philosophers have allowed epistemic appraisal to eclipse heuristic appraisal, allowing epistemic status to take credit for what are really heuristic features of theories, models, and experimental results.

Here is a list of some (overlapping) ways in which scientists deal with abrupt representational change, with some added commentary. Some of these items will further articulate scientists' differences from philosophers.

1. Creative scientists are quick learners who adapt readily to changing circumstances, especially in their early years. After all, many scientists enjoy major switches of research interests without this being forced upon them.
2. While individual scientists can be adaptive (practices can go obsolete even when their former practitioners do not), scientific communities can be adaptive in a trans-generational sense, which, in extreme cases, produces the so-called Planck effect of the older generation simply dying out without being converted to the new ideas and practices. In progressive fields, the teaching methods and technical workshops respond rapidly to the changing scene. Recall the summer seminars on hot biological topics at Cold Spring Harbor.
3. Scientists are risk-taking, pragmatic opportunists who go where the action is, or promises to be. They can respond quickly to research and funding contingencies, e.g., by altering their project goals and methods. In this respect they can be remarkably agile rhetorically – "unprincipled" pragmatists. For pragmatists, success is its own justification. To construct something that works is their goal, whether or not it is the goal their project originally targeted. Pragmatic scientists are more interested in working models and mechanisms than in metaphysics.
4. And there is more than one way of working.[9] Scientific advance can come in several forms and dimensions, so defects in one dimension are not fatal. A defective hypothesis may yield problem solutions or empirical results too interesting to ignore, e.g., a conceptual or technical breakthrough, a "how possibly" explanation, or a confirmed novel prediction.[10] Pragmatists do not worry overly much about

[9] In this regard, Francis Bacon's experimental philosophy was perhaps the first articulation of a pragmatic attitude toward scientific results. Previously, "Greek" natural philosophy was a highly intellectual enterprise. Bacon's view (on my interpretation) was that if you can do X to nature and get nature reliably to respond Y, this is an important result *whether or not* you have yet achieved full intellectual understanding of the process. This was a significant departure from the Aristotelian worry about intervening in nature causing artifacts.

[10] There is a controversy in confirmation theory over whether novel predictions have extra probative weight within the accounting scheme of epistemic appraisal. I side with Mill, Keynes, Laudan, and Brush (against Herschel, Whewell, the Popperians, and the Lakatosians) in denying that they do. However, novel predictions *are* certainly of special interest from the standpoint of heuristic appraisal. They give scientists new things to investigate and, in some cases open up whole new domains.

interpretation as long as they can get problem solutions.[11] But progress can also consist in providing theoretical or mathematical grounding for a hypothesis (such as Planck's empirical radiation formula) or a questionable technique (such as the use of the Dirac delta function). Empirical well-foundedness can occur in the face of conceptual difficulty.[12]

5. Working scientists are, and must be, presentists and perhaps also whigs.[13] They interpret past work in terms of their own present problems, sometimes deliberately, sometimes unintentionally. That is, they deform previously solved problems so as to achieve a close enough match with their current problems, a point that Kuhn (1970, "Postscript") makes about normal science and the genealogies of problems that it produces. Good science makes for bad history (Nickles, 1992, 2003a). This point includes rewriting the past history of the discipline to locate, celebrate, and inspire present work.[14]
6. Item 5 sometimes involves the literary device of "strong reading" (Bloom, 1973). According to Kuhn (1978), it was really Einstein and Ehrenfest who invented the early quantum theory in 1905 by mistakenly attributing to Planck a solution that he in fact rejected to a problem that he did not have (the problem that Ehrenfest dubbed

[11] Nancy Cartwright (1987) confirms the suspicion that American quantum theorists did not regard themselves as culture carriers to the extent that European scientists did. Solving problems was what really mattered, not providing cultural interpretation of the results. Few physicists and college professors today can offer a very coherent explanation of what underlying physical processes are occurring when they formulate and solve their Schrödinger equations. And the work of those physicists who did worry, and tried to do something about it, including Einstein and Bohm, was not taken very seriously as physics. But what about Bohr? He was a culture carrier and worried deeply about the interpretation of quantum mechanics. Did not his Copenhagen interpretation help to advance the field? Without claiming to be an expert in these matters, I suggest that the answer is yes and no. Many physicists found Bohr's more philosophical writings to be obscure. And most of those who were helped by the Copenhagen interpretation would probably say that the reason was pragmatic. To them Bohr's message was: "While we must still use classical language in setting up our experiments, there is no way to make intuitive sense of quantum processes in classical language. So just get on with your work and don't worry about it!" (These are my words, not Bohr's – TN.) The kind of physical intelligibility sought by Kelvin in his famous remarks concerning the necessity of mechanical models to scientific understanding, and also the kind of intelligibility still sought by Einstein, are not to be had – although Einstein's emphasis on theories of principle versus matter theories may have encouraged the later de-emphasis of causal mechanisms.

[12] For amplification of this point, see Laudan (1984, Chap. 4), who criticizes Kuhn's holistic claim that scientific revolutions change everything at once: theoretical claims, aims and standards, methods and experimental practices.

[13] Depending on how whiggism is formulated, it overlaps presentism, more or less. Whiggism as Butterfield (1931) characterizes it covers multiple sins, only one of which is history being rewritten from the point of view of the victors. See Hull (1989, Chap. 12, "In Defense of Presentism").

[14] There is something genetic (history-based) about intelligibility. As Stephen Toulmin remarks (in praise of John Dewey):

> "knowing" is intelligible only as the outcome of the activities by which we "come to know" the things we do; and all the different ways of "coming to know" involve corresponding methods of *operating on* the world. (1988, p. xi, Toulmin's emphasis)

In ignoring context of discovery, the logical empiricists therefore put all their intelligibility eggs in the basket of logical structure, leaving no role for historical motivation. How, then, can they explain why technical scientific articles often contain a historical section locating and motivating the results being presented? Of course, this scientific history is usually reconstructed, "presentist" history, but that just makes its function of enhancing the intelligibility of the present work all the more obvious. Such articles often contain some "heuristic appraisal" comments toward the end in which it is indicated where the presented work may well lead. Thus the work is given a possible future as well, which is intended to enhance both its intelligibility and its perceived importance.

"the ultraviolet catastrophe"). According to Kuhn, Planck was still working within the classical statistical mechanics of Clausius and Boltzmann. But perhaps Planck's paper suggested meanings that he did not explicitly intend.[15]

7. Many Kuhnian exemplars (standardized problems and solutions that guide research) usually survive a major representational change, or can be adapted to the new research context. Thus the Bohr–Sommerfeld approach to the old quantum theory could start from classical formulations and then impose quantum conditions. Ehrenfest's "adiabatic principle" enabled physicists to extend this approach to new mechanical systems. Later, "Ehrenfest's theorem" permitted many quantum mechanical problems to be set up using the resources of classical, Hamiltonian mechanics. This point is even more obvious for experimental research, both results and techniques.
8. Scientists freely employ rhetorical tropes – especially analogy, similarity, and metaphor – to accomplish the above. They do not generally limit themselves to static logical or semantic schemes, although rigid operationism was a temporary exception.[16] From Plato, Descartes, and the founders of the Royal Society to the logical empiricists, philosophers have preferred logic and literal, "plain" language

[15] In attempting to debunk as a myth Planck's being the founder of quantum theory, Kuhn's history is controversial. I personally think it is quite plausible. There is little evidence that Planck really intended to introduce quantized physical processes even of emission or absorption in 1900. His introduction of the quantum of action was little more than a mathematical trick, made in desperation, as he admitted. Nonetheless, I think Kuhn is unfair to Planck here. Kuhn's almost psychoanalytic approach of trying to get inside people's heads led him to identify the physical meaning of Planck's work with Planck's own conscious intentions. Literary theorists long ago gave up "the intentional fallacy" of attributing complete and final authority of the meaning of a work to its author. Some go so far as to speak of the disappearance of the author.

Scientists provide strong readings all the time in an effort to distill out the most basic meaning of central claims. Historically, Planck's empirical law was theoretically analyzed and rederived several times, by Ehrenfest, Debye, and others, including Einstein on several occasions. What makes a work classic is that it provides the basis for a seemingly endless number of strong readings. Moreover, some of the credit does redound to the author, who, after all, did write a work rich enough to sustain such interpretations. So in my opinion Kuhn should have interpreted Einstein and Ehrenfest as giving a strong reading of Planck.

In fact, Kuhn's treatment of Planck is rather surprising by his own lights, for he is careful to deny that the fullness of communication and unity of agreement among normal scientists is a matter of a consensus of private beliefs. On the contrary, if asked to explain their reasons for doing something one way rather than another, they might well disagree. That is to say, much of their know-how is subarticulate. Thus there can be quite a bit of slack between scientific behavior and private beliefs. On the other hand, this position does not square well with Kuhn's internalist psycho-history methodology.

[16] In 1927, at the beginning of *The Logic of Modern Physics*, P. W. Bridgman (later a Nobel Laureate) wrote:

> We should now make it our business to understand so thoroughly the character of our permanent mental relations to nature that another change in our attitude, such as that due to Einstein, shall be forever impossible. It was perhaps excusable that a revolution in mental attitude should occur once, because after all physics is a young science, and physicists have been very busy, but it would certainly be a reproach if such a revolution should ever prove necessary again. (1927, p. 2)

Had Newtonian physicists defined the concept of simultaneity more carefully, Bridgman said, no Einsteinean revolution would have been necessary. On this view, all future conceptual needs and refinements can be known in advance of theory and evidence, and future theories can be formulated in a fixed conceptual language already present! Kuhn's and Feyerabend's views were diametrically opposite: revolutions are necessary to break out of the old conceptual framework. Shapere and Toulmin rightly complain that both positions depend upon untenable views about meaning. See Toulmin (1972), Shapere (1984, Chaps. 3, 5). Operationism as a rigid philosophical doctrine is surely one of those obstacles about which Bachelard would complain.

to rhetoric; however, the growth of language and its expressive power depends heavily upon rhetoric. Another reason for the neglect of rhetorical devices is philosophers' excessive concern with epistemic justification. Arguments from analogy and metaphor are weak, after all. But they can be helpful indeed in context of discovery and in heuristic appraisal. Mary Hesse (1965), Feyerabend, and Kuhn introduced rhetoric into philosophy of science during the 1960s; and for all three it is utterly essential to the scientific enterprise.

9. Despite their concern for precision, scientists can differ markedly in their personal scientific beliefs and can be historically casual in their retention of received terminology, relative to the demands of some philosophical theories of meaning and reference. Scientists want their own language to be retained even if the interpretation differs.[17] That way, they still get credit. They are implicitly banking on presentism. (Consider the history of the terms "oxygen," "atom," "electron," and "gene"). And physicists can have sharp verbal disagreements about what exactly quantum theory asserts about reality while agreeing completely on how to work the problems. In the mature sciences, unlike philosophy, such verbal or fideistic disagreement need not be the occasion for breakdown. It is fortunate that attempts to make scientific meaning completely precise have failed, for a coarse-grained approach to meaning is essential to scientific creativity (or so I believe). It guarantees a certain amount of private and public interpretive variation around claims put forward, and occasionally a variant proves superior to the original. To turn the point around, virtually no major scientific claim has the "meaning" that it had when originally proposed, say 20 or 30 years before.
10. In general, scientists are not hampered in the way traditional philosophies of language predict that they should be. Donald Davidson (1986) makes the general claim that, really, no one is hampered in this way. We all manage our linguistic acts of transmitting and receiving "on the fly," as it were, and for the most part successfully. According to standard philosophical and linguistic theories, such communication should be impossible. But surely successful innovative practices in science and elsewhere ought to constrain and test our theories of language rather than vice versa.
11. As both the Bohr story and Kuhn's discussion of science in crisis suggest, there are times when a scientific community is especially open to radical new ideas. During these times anything that hints at a breakthrough will be worth examining. So many strange developments have happened in twentieth-century physics that physicists today may be more open than those of the late nineteenth century, whether or not they perceive their subject to be in crisis. At certain periods, wider cultural developments such as modernism in the arts or Weimar culture may have reinforced the quest for novelty (Forman, 1971).
12. One way in which scientists can identify which new work is relevant to their own is to retreat to experimental determination of key parameters. For example, if Bohr's electron has the same charge and mass as Lorentz's electron, that is one good reason for recognizing Lorentz's work and Bohr's work as pertaining to the same subject (Arabatzis, 2006).
13. Creative scientists excel at intellectual play.

[17] David Hull once made this observation in discussion.

In summary, I reiterate my point that scientists have a lot of experience in dealing with semi-coherent materials, as we see when we pay closer attention to their practices at the frontier of research, that is, in the "context of discovery," broadly speaking. Initial ideas and practices almost always possess major defects, including departures from orthodoxy. This is not a shocking situation for scientists.

Beyond these sketchy remarks, I shall not try to justify my negative claim – that Kuhnian incommensurability is not as serious a problem for scientists themselves, as Kuhn himself and other philosophical commentators make it out to be. Rather, in Sect. 5 I shall offer some alternative models of disruptive transformation. First, however, it will be helpful to expand on some Kuhnian themes.

4. KUHNIAN REVOLUTIONS: REPRESENTATIONAL OR PRACTICE-BASED?

I claim that it is important to distinguish theoretical-representational ruptures from disruptions of practice. The former are not always the cause of the latter. In fact, in some cases the direction of the causal arrow is reversed.[18]

Kuhn always said that he thought his account of scientific revolutions was his most original contribution (e.g., Kuhn, 2000). The *locus classicus* is the famous (or notorious) Sect. X of *Structure*, "Revolutions as Changes of World View." As the title suggests, Kuhn unpacked much of the idea in terms of visual perception and visual metaphors – Gestalt switches and the like (an emphasis that he later regretted). Many of his examples involved scientific observation (astronomical and experimental) and psychological studies of perception. He began with rather pure observation, then introduced perceptual shifts occasioned by changes in paradigm (all of which was a visual preview of what he would later say about the learned similarity relation in the "Postscript" and other essays). He attacked the "Cartesian" myth of the given, according to which perception is the given plus theoretical interpretation. On Kuhn's view of that time, theory change causes meaning change and perceptual change all the way down, resulting in a disruption of practice (Shapere, 1984, Part I; 2001). Kuhn considered this his major contribution to the discussion of theory-ladenness of observation, meaning change, and cumulativity – themes that, during the 1960s, lay at the heart of a serious reassessment of the epistemology of empiricism.

This material plus Kuhn's leading examples of scientific revolutions – the Copernican Revolution (or the entire Scientific Revolution through Newton), the chemical revolution, relativity, and early quantum mechanics – naturally suggested to interpreters that Kuhnian revolutions either found a new field (the onset of the first paradigm) or involve a paradigm switch within a mature field, and that incommensurability

[18] Andy Pickering's early writings on the role of opportunity and fertility assessment on research choices in high energy physics have especially influenced my thinking about heuristic appraisal. As an old pragmatist myself, I appreciate his emphasis on pragmatic opportunism in scientific practice. See, e.g., Pickering (1980a, b, 1984). See Léna Soler's chapter in this volume for a detailed account of Pickering and Ian Hacking on incommensurability. Recall also Laudan (1984) on breaking up the big Kuhnian package.

(perceptual and conceptual displacement, living in a different scientific world) is the hallmark of revolution.

On the representational construal of Kuhn, revolution, as a revolt *against* something, turns on a basic rejection of established doctrine, the overturning of previously articulated “dogma.” A revolution, understood thus literally, is a discontinuity of this special kind. The more general terms “rupture,” “disruption,” and “discontinuity” allow for other kinds of breaks. Of course, today, thanks largely to the advertising industry, it has become common to use “revolution” and “revolutionary” to mean any supposedly significant advance.

As indicated, virtually all of Kuhn’s examples came from pre-World War I physics and early nineteenth-century chemistry. A greater variety of examples would have challenged this one-track conception of revolution. Yet even his own examples do not fit the above description neatly. Einstein’s special theory of relativity, for example, though limited in one sense (in its application only to inertial systems), was far more than a disruption contained within classical mechanics. For one thing, it also deeply involved electromagnetic theory. For another, Einstein’s two principles – the principle of the constancy of the velocity of light and the relativity principle – were advanced as general constraints on all of science. There can be no instantaneous action at a distance, and the laws of nature must be expressible in the same form in all inertial frames. Third, more than anyone else, Einstein’s example made purely theoretical physics respectable. And his “theories of principle” approach, in contrast to the matter-theoretic “constructive theories” of H. A. Lorentz and others, brought a new style to physics. Fifth, Einstein’s relativity theory also elevated theory of measurement from a low-level topic to one of high scientific interest – one that would, of course, become much higher still in the mature quantum theory. Kuhn’s (1962) focus on meaning change as the basis for the revolutionary break fails to track these more important aspects of the change in direction that physics took from Einstein on. Indeed, in my opinion Kuhn’s large-scale distinction of normal from revolutionary sciences fails to capture the still larger-scale transformational changes in physics that began in the nineteenth century, in “classical” physics, the sweeping changes that brought modern mathematical physics into being. Another sweeping transformation was that of the general, scientifically informed, world view. We must not forget that many late nineteenth-century physical scientists were convinced that reality was within their theoretical and experimental grasp, and that that reality was the deterministic mechanical lawfulness of ordinary-sized, observable bodies. The radical new physics of the unobservably small and the incomprehensibly large would soon dash this common-sense world picture much as the Scientific Revolution had dashed the commonsense world of Aristotle.

Please note that I am not totally denying the existence of incommensurability in Kuhn’s original sense, whatever it is. But I believe that it is more a problem for philosophers and internalist historians than for working scientists. Kuhn’s later, restricted sense of incommensurability may be more defensible within cognitive theory, but it no longer has the radical philosophical implications that Kuhn advertised in the 1960s, as far as I can see.[19]

[19] See Barker, et al. (2003) for an excellent account of Kuhn’s final work on incommensurability.

Citing such passages as Pauli's lament that he had no idea where physics in the mid-1920s was going and that he wished he had been a shoe salesman, my critic may object that I am being too hard on Kuhn. It is true that the break with classical theory was disconcerting, but I suggest that what really bothered Pauli and others like him was that it was no longer clear what the good problems were or how to attack them. "But," my critic replies, "that is precisely because of the breakdown in theory." Well, yes and no. To say that the known falsity of classical theory is a sufficient explanation makes practice too dependent on true (or believed-true) theory. What the critic misses is that theories and models can have great *heuristic* value as practical guides. The breakdown of the classical theory showed that the heuristic intuitions it afforded were no longer reliable. It is not the epistemic breakdown or the semantic break in itself that is the problem; rather, it is the *heuristic breakdown*, the alienation of theory from practice. Here is one instance in which philosophers have allowed epistemic appraisal and the resulting epistemic status to take credit for what are really heuristic features of a theory or model.

If we now return to our question: To what degree is *Kuhnian (representational) incommensurability* a major source of recognition problems, what should we answer? Given the standard, philosophical answer inspired by Kuhn's original work, that theory change drives meaning change, perceptual change, changes in goals and standards, and the disruption of practice, my response is "Incommensurability is overrated." Deep theory change is neither necessary nor sufficient for disruption of practice. As Laudan (1984) argues, we should not tie everything together in a single package.

If that is so, we need to seek alternative accounts of the kinds of deep scientific change that actually occur. In Sect. 5 I shall consider alternative models of radical change authored by Richard Foster, Sarah Kaplan, and Clayton Christensen for the business world, models that, suitably adapted to science, may fit cases of the kind that Kuhn's model ignores. But first let us note Joseph Rouse's alternative interpretation of Kuhn himself, an interpretation that emphasizes scientific practices over representational theories.

Rouse (2003) provides a generous reading of Kuhn that makes Kuhn's position more flexible and up to date and that anticipates some features of the Christensen models as well as the forward-looking, pragmatic impetus that I ascribed to working scientists above. Rouse proposes a reading of Kuhn based, in Kuhn's words, on "the research activity itself," that is, scientific practices. On this view, Kuhn's central interest is not in epistemology, certainly not in the sense of traditional, backward-looking confirmation theory, but in showing how science works. Paradigms are not objects of belief, and the associated generalizations do not constitute a credo. Rather, they define the field of scientific practices, the field of possible action, for a given specialty and the research opportunities that it affords. Writes Rouse,

> Accepting a paradigm is more like acquiring and using a set of skills than it is like understanding and believing a statement.… Scientists *use* paradigms rather than believing them.… Kuhn has often been misread as insisting that members of scientific communities are in substantial agreement about fundamental issues in their field. What he actually says is that normal science rarely engages in controversy about such fundamentals. A lack of controversy is quite consistent with extensive disagreement, however, if research

> can proceed coherently and intelligibly without having to resolve the disagreements. Shared paradigms enable scientists to *identify* what has or has not already been done, and what is worth doing, without having to agree on how to describe, just what those achievements and projects are. (2003, pp. 107–108, 110, Rouse's emphasis)

I find this position (which Rouse himself endorses) congenial. Whether or not it is the best interpretation of Kuhn rather than a sympathetic backward look I shall leave to experts such as Paul Hoyningen.[20] I agree with Rouse that the disruption of practical skills, including an expert ability to appraise *heuristically* the field of action, provides a better explanation of many ruptures than Kuhnian incommensurability as standardly interpreted.

From this viewpoint we can see that practices provide a broader basis for considering disruptive change than do big, representational theories. There can even be disruptive change where no major theory clash is at stake, just as we can envision major theory clashes that require no equally great disruption of practices. (Special relativity is surely a case of the latter sort, once scientists got over the difficulty of recognizing Einstein's strange-looking 1905 paper as a serious contribution to physics.) Even when theory clashes, incoherencies, or translation failures cause grief, it is not because of them *per se* but because of their disruption of scientific practice (including theoretical problem-solving practices). Since Kuhn does sometimes recognize the importance of these points, there is an unresolved tension, either in Kuhn himself or in the competing interpretations of his "real" position. Gerd Gigerenzer (2000) adds the important point that deep theoretical innovation itself is often inspired by prior practices, which will, in turn, alter the centrality of those practices. I'll return to this point below under the head of "low-end disruptions."

If Rouse's is a defensible interpretation, then Kuhn, in effect, turned their own discovery-justification distinction *against* the logical empiricists and Popper. The important thing for philosophers to do (he said, on this reading) is to join historians and sociologists in figuring out *how science works*, especially at the frontier of research – to understand the very practices that fall within the excluded category of context of discovery. And much of what the philosophers say about context of justification is almost useless for understanding real science. As Kuhn once put it, looking at the final products of scientific work as published in summary papers and textbooks is like trying to understand the lifeways of a foreign country from a tourist brochure! On this reading, Kuhn anticipated the turn away from theory-centered accounts of science and the projects of representation and justification. He turned toward the postmodern project of understanding practices.[21] And on this reading, "change of world" just means a transformation of the life of research, which may be occasioned by a new instrument or technique or by an unanticipated discovery such as Roentgen's discovery of x-rays[22] as well as by a new theoretical contribution.

[20] Hoyningen (2006) reports that the later Kuhn regarded his comment at *Structure* (1970, pp. 8 ff.) about transcending the discovery-justification distinction as a "throwaway remark" of small importance. But if Rouse is right (and if we do not privilege the later Kuhn's take on his earlier work), we can read *Structure* as bearing strongly against the standard application of that distinction.

[21] While I agree with much that Rouse says, I am uneasy about his proposal simply to fold philosophy of science into culture studies (Rouse, 1996, Chap. 9).

[22] For Roentgen, see Kuhn (1977, Chap. 7), "The Historical Structure of Scientific Discovery."

5. SOME NON-KUHNIAN (?) VARIETIES OF DISRUPTIVE TRANSFORMATION

5.1. Organizational transformation

It has long been recognized that Kuhn's account neglects the sort of revolutionary transformation that several sciences have undergone during the twentieth century, especially since 1940. Here we find the vast reorganization of scientific research that Derek Price (1963) dubbed the change from "little science" to "big science." This has been disruptive in many fields, where traditional, low-cost "table top" experimentation has been shoved to the periphery of scientific practice. There are many kinds and degrees of big science, of course, with several varieties of institutional support. Big science need not involve a national laboratory with a huge accelerator. We might also speak of "medium-sized science" such as that which requires access to a modern biochemical or molecular genetics laboratory or access to supercomputing of some kind. In many areas, the varied technical expertise required to do good experimental work requires the collaboration of many colleagues across several fields. This degree of cooperation was unusual in the past and requires new kinds of skills to manage.

5.2. Disruptive technologies

Kuhn's early papers on Sadi Carnot indicate his awareness that major innovation may come from the outside, from another field, or even from outside mainstream science as a whole as it then exists. Carnot's fundamental work on the heat engine seemed, even to historians, to drop out of the sky – until Kuhn (1955, 1958, 1961) discovered that Carnot was working in a specific engineering tradition. But Kuhn seems to have forgotten this lesson in the wake of the overly rigid structure of *Structure*.

Recent publications on capitalistic business enterprise suggest some other patterns of disruption. Business analysts Richard Foster and Sarah Kaplan (Foster, 1986; Foster and Kaplan, 2003) contend that business life has accelerated during the twentieth century as the pace of innovation has increased, so that the average lifespan, even of good companies, is becoming shorter. Why? Because "the market" tracks innovative success better than most any corporate enterprise can. Legally speaking, a corporation can live forever, but in practice even successful companies fall victim to the conservative mindset of their directors, according to which they must protect their current product leaders and market shares. The "wrong" sort of innovation would undermine their own best-selling products or services. They become wedded to what Foster and Kaplan term "the continuity assumption," which, in addition to familiar expertise, may include a sentimental attachment to their traditional business lines and customer base. Foster and Kaplan contend, by contrast, that only those who see business life in Joseph Schumpeter's way, as a form of "creative destruction," are fully prepared to be pragmatic opportunists. They distinguish three rather than two basic types of innovation: incremental innovation, substantive innovation (which can involve sudden major advances within a product line or paradigm), and transformational innovation. Only the last is truly "competency destroying" (p. 113).

Given that basic science, too, is driven by intensely competitive innovation, we might ask whether there has been a similar acceleration in twentieth-century science, with the growth of big science somewhat paralleling that of big business. (I am talking about basic science in general, not only areas such as biomedical research that are closely wedded to industry.) Kuhn stressed the conservatism of normal scientists, and he is frequently cited by business analysts on just this point.[23]

Another business analyst, Clayton Christensen (1997, 2003) of the Harvard Business School, distinguishes what he terms *sustaining technologies* from *disruptive technologies*. The former can include major, even revolutionary, innovations, as when automatic transmissions were introduced to automobiles and jet engines to aircraft.[24] In business terms a sustaining technology is a successful technology that attempts to keep its established customer base happy by introducing improvements to popular product lines. Indeed, this seems to be exactly what a successful company and its executives should do. Christensen was originally led to his investigation by asking the question why such successful and well-managed companies sometimes fail – and fail precisely *because* they pursued the "obviously successful" strategy. This phenomenon is vaguely reminiscent of Kuhn's claim that successful normal science leads to revolution, but there are important differences.

On Christensen's view, a big technical breakthrough (technical revolution) is neither sufficient nor necessary for major discontinuities in business life. As I see it, the scientific analogue is my claim that major technical innovation (whether theoretical or experimental) is neither sufficient nor necessary for disruption of practices. Another observation is this. While Kuhn deserves credit for pointing out that much of the debate between a paradigm and its challengers concerns their heuristic potential as a guide to future research (rather than simply their confirmation track record to date), he talks as though it takes a crisis to stimulate essential innovation. But in the business world, if Christensen is correct, the main threat comes from developments long underway outside the core of the business paradigm. In science, the Carnot case comes to mind, and I shall give some other examples below. But first let me elaborate Christensen's view.

By contrast with a sustaining technology, a disruptive technology begins as a fringe phenomenon that initially poses little threat to the aforementioned customer base, for the technology and accompanying products and services are either inferior to those currently dominant or irrelevant to current market shares. Such a technology may begin among tinkerers and amateurs, but as it grows and prospers it threatens to displace the high-end products of the first kind and may eventually swamp the successful companies themselves. A crucial point is that these technologies need involve no major technical breakthroughs, no revolutionary theoretical developments in order to get underway.

23 This in turn raises science policy questions, given the degree to which science is now supported by government and private industry. Have the sciences and their funding agencies become institutionally conservative to a degree that will hinder future innovation? We might then expect to find a few scientists analogous to those business entrepreneurs who attempt to break out of this innovation boom-bust cycle as paradigm change ossifies into business-as-usual.

24 Recall that Imre Lakatos' "methodology of scientific research programs" and Larry Laudan's research traditions already allow for major innovative advances *within* the established tradition (Lakatos, 1970; Laudan, 1977).

Christensen (2003, Chap. 2) proceeds to distinguish two kinds of disruptive technologies, which can occur separately or in combination. As the name suggests, *new market disruptions* appeal to a previously non-existing market, so such technologies, initially, do not compete for established markets. Rather, they compete with "nonconsumption" (p. 45). One example is the Sony Walkman, which did not originally compete with household radios and stereo systems. Another is the personal computer. Both of these products created new markets where none existed before. In the USA the Apple was the first highly popular personal computer, but most members of the scientific and business communities considered it a toy, a household gadget, or an educational tool for school children. Eventually, as we know, both the miniature audio system and microcomputer industries became so successful as to undermine established markets for large installations. A third example is desktop copiers from Canon that made photocopying easy for individuals in their own offices rather than taking documents to a company center of large machines run by technicians. The ease of copying resulted in a tremendous increase in the quantity of copies made, thus expanding the printer market and making a joke of the idea of the paperless workplace. A much earlier example was Kodak's introduction of simple "Brownie" cameras that anyone can use. The user-friendly "point and shoot" feature is still promoted in advertising campaigns today. This development put many professional photographers out of business.

Christensen's other type of disruptive technology is *low-market* or *low-end disruptions*, which "attack the least-profitable and most overserved customers at the low end of the original value network" (p. 45). "Overserved" customers are those who don't need the fancy bells and whistles – and higher cost – of the high-end products. Christensen's leading examples are discount airlines such as Southwest, discount retailers such as Kmart and Wal-Mart in America, and small, specialized Japanese steel mills that initially made low-grade products such as rebar using lower-cost smelting techniques than the big companies with their huge blast furnaces.

Let's take this last example a step further. As their manufacturing processes improved, the small mills produced higher and higher quality products, more cheaply than the big mills such as US Steel and Bethlehem Steel. Eventually, Japanese mills were able to undercut the established markets and drive some of the biggest mills out of business, completely changing the face of the industry. Naturally, this collapse threw thousands of skilled people out of work once their practices were seen to be outmoded. Similarly, the discount retailers, once they gained a foothold, moved more and more upscale, gradually cutting into the established markets of the big department stores and expensive specialty stores. The Wal-Mart phenomenon has ruined hundreds of local business districts across America and undercut large companies such as Sears.

Some of these examples are not pure, as Christensen points out. Southwest Airlines, for example, is a hybrid. It created new markets, starting in the southwestern USA, by encouraging people to fly between cities that are nearby by western US standards but distant by eastern standards; but it also captured the low-end, least-profitable routes of big airlines, such as American Airlines. The same could surely be said for European carriers such as Ryan Air.

This is a simplified account of Christensen's models, but what strikes me about the sustaining and disruptive models already is that they do not fit Kuhn's pattern well.

On the one hand, major innovative spurts can occur as a fairly normal result of research and development within established product lines. Great leaps forward need not produce major disruption. On the other hand, a truly disruptive displacement of Christensen's sort need not have its source in a revolutionary technical breakthrough. Revolutionary developments are neither necessary nor sufficient for disruption.

Are there scientific analogues to Christensen's models of disruption and non-disruption?

6. WHAT MIGHT DISRUPTIVE SCIENTIFIC DEVELOPMENTS LOOK LIKE?

First, it is worth noting that our knowledge of history gives us a skewed view. Many of the developments that Kuhn himself discusses, for example, began as the scientific counterparts to something like new-market and/or low-end disruptions. This point is virtually entailed by his own account, according to which normal science consists in the thorough investigation of esoteric details and scientific revolutions are nonlinear responses to seemingly ordinary anomalies that simply cannot be resolved, no matter how much attention they receive.

This is true of the principal experimental routes to quantum theory. Spectrographic analysis began as a new-market science relevant to chemistry but not to mechanics. Measurements of specific heats initially belonged to thermodynamics, not mechanics. The photoelectric effect was initially an esoteric curiosity. Kuhn's own account shows how the problem of blackbody radiation arose out of researches sponsored by the German power industry and that Planck thought he could achieve a classical solution to the problem of finding and explaining the energy distribution formula. Initially, Brownian motion was a curiosity for biologists. It was not until Einstein showed its relevance to the atomic-molecular theory that it became important. In all of these cases, and many, many others, it is only with whiggish hindsight that we regard them as fundamental from the beginning!

Second, when we look to the commercial production and distribution of scientific instruments and such research resources as chemicals, model organisms, and assay services, we find a direct connection between the business world and basic scientific research (yet still not the sort of direct commercial ties characteristic of the pharmaceutical industry). Technological change can drive rapid scientific change in a given domain. For example, in laboratories, many manual skills have been replaced by automated procedures. Specialized equipment of other kinds has altered scientific practices enormously. New-market technologies have stimulated the emergence of corresponding new market research techniques and entire scientific specialties. Think of scanning and imaging technologies. Think of all the devices and procedures now available in a fully equipped bioscience institute. Think of computers.

Initially, computers were relatively low-end devices employed to relieve the routine of tedious calculations, inventory tracking, payroll sorting, and the like. The early ideas of Charles Babbage and Ada Lovelace had been inspired by the mechanized Jacquard looms, with their punch cards, in the British textile industry. Since about 1950, as their power and speed have increased, they have moved "up market," scientifically speaking, transforming some fields and creating others. Whole fields of computer science have emerged to support scientific and commercial enterprises. Nonlinear dynamics and complexity theory could hardly emerge as a field until fast computers made them possible. Poincaré had to perform his early work on what we call "deterministic chaos" by hand. High-speed computers are essential to modeling across many fields, and, increasingly, computer simulations are a useful substitute for nature in certain kinds of experimental work. In recent years the remarkable emergence of evolutionary computation (at first just a kind of esoteric, even crazy, idea comparable with quantum computing today) has provided a whole new approach to problem solving. Evolutionary computation is now entering the domain of genuine creativity.

A different sort of candidate for new-market and subsequently a low-end disruption in the sciences is statistical-probabilistic techniques. Probability theory, as we know, originated in gaming questions; and statistics in the dull social-demographic records of birth, marriage, death, and the like (Hacking, 1992). Gauss and others subsequently developed an elementary theory of error to handle the noise of variant observations, while Laplace saw probability cum statistics as a way of extracting knowledge from ignorance, or higher-grade knowledge from lower-grade. (It would later grow into sophisticated treatments of experimental design.) This was still pretty low-end stuff, designed more to protect current theory against wayward observation than to challenge it. But then, in the 1850s and 1860s, there was a flashover into serious physics, in the kinetic theory of gases, which soon grew into full-blown statistical mechanics. Hacking speaks of a whole new "style of reasoning" coming into being, something larger even than a Kuhnian paradigm. That probability and statistics now lay at the heart of physical theory was a major step toward modern, mathematical physics. No doubt this step left many old-style physicists unable to keep up. The practices of physicists changed enormously. Maxwell and Boltzmann finally concluded that the "statistical" of statistical mechanics was essential in that the law of entropy was irreducibly statistical, although many nineteenth-century thinkers demurred, contending that the statistics only reflected human ignorance. The new way of thinking would gain ground rapidly after Bohr's atom models of 1913, so that, by the 1930s, experts such as John von Neumann (1932) could argue that quantum mechanics is logically incompatible with deterministic "hidden variable" models. By then nature itself was considered inherently probabilistic. Probability and statistics had risen from mere political arithmetic and error analysis to constitute the very core of mathematical physics! On a smaller scale (so far), something similar may be happening today in nonlinear dynamics (Kellert, 1993).

Gerd Gigerenzer (2000 and elsewhere) extends the story to the cognitive sciences with his thesis that research tools are often the source of *theoretical* innovations. Statistical-probabilistic thinking has transformed research design in all scientific fields to the extent that experimental analysis not conforming to its canons is immediately called into question. But in the cognitive sciences these inferential tools came to suggest new models for how human cognition itself works, e.g., human beings as Fisherian or Bayesian agents. A similar story can obviously be told about the digital computer, and computer models of the mind (cognition), which, since the 1960s, has transformed psychology into modern cognitive science.

After all, it is a familiar idea to scholars that historical peoples have always tended to model the unknown in terms of their latest technologies; and we are no different. Once the idea is broached, we can think of many other examples. Among the many that Gigerenzer himself notes are clocks and other mechanical devices that caught the attention of early natural philosophers and were sometimes used as scientific instruments. Before long Descartes and others had turned nature itself into clockwork and animal and human bodies into machines, thereby not only mechanizing the world picture (Dijksterhuis, 1961) but also completely transforming how these topics were investigated. During this same period, natural philosophers were struggling to employ arithmetic and algebraic techniques in the description and explanation of natural phenomena. The Aristotelian tradition had regarded mathematics as a distinct subject, inferior to natural philosophy, and the new arithmetic had been used mostly for keeping business accounts. Galileo, Newton and company radically altered the practices of natural philosophers. Ingenious though he was, Robert Hooke suffered from the fact that he was not a mathematician. The transformative mathematization of new fields continues to this day. A current example is ecology (Kingsland, 1995).

Notice that Gigerenzer's thesis reverses the relationship between theory and practice assumed by many philosophers and internalist historians. In the cases he considers, practice inspires theory rather than vice versa. Of course, things can go in either direction.

In recent decades the biosciences have surely undergone many disruptive changes with new experimental practices and models rapidly rendering the previous winners obsolete. Less than two decades after Watson and Crick, Gunther Stent could write in his 1971 textbook:

> How times have changed! Molecular genetics has … grown from the esoteric specialty of a small, tightly knit vanguard into an elephantine academic discipline whose basic doctrines today form part of the primary school science curriculum. (1971, p. ix)

Stent had been a member of Salvatore Luria and Max Delbrück's Phage Group. Recalling the Delbrück story will illustrate both the disruptive-transformative change that can be introduced by immigrants from other fields but also how normalized science can drive creative people out of an established field into unstable, unexplored territory.

Inspired by Bohr's 1932 lecture, "Light and Life," and Schrödinger's little book, *What Is Life?* (1945), several physicists, including Delbrück and (later) Francis Crick, switched to biology in order to study the problem of gene replication and stable inheritance, because that seemed to be the most interesting yet now perhaps tractable "big problem" left. Via the American Phage Group, founded by Delbrück and Luria, and a British group taking a structural approach to molecular structure by means of x-ray analysis, molecular genetics emerged as a new field with a style closer to physics and physical chemistry than to traditional biology. Delbrück and the Phage group found traditional "blood and guts" biology revolting. Delbrück notoriously disliked even biochemistry and anything having to do with medical applications, as diverting science from the really fundamental problem of genetic multiplication, which he expected to have the features of a deep paradox that might require introducing new laws of nature. Reminding his audience that the Bohr atom model totally abandoned the intuitively reasonable classical mechanical model, according to which there should have been mechanical vibrations within an atom corresponding to each of its spectral frequencies, Delbrück (1949) drew an analogy to biochemistry. Plausible as it may seem that a cell is just a

> [S]ack full of enzymes acting on substrates converting them through variations intermediate stages … this program of explaining the simple through the complex smacks suspiciously of the program of explaining atoms in terms of complex mechanical models. It looks sane until the paradoxes crop up and come into sharper focus. (1949, pp. 18, 22)

Delbrück hit on bacteriophages as his experimental item of choice not only because they multiplied in only half an hour but also because they were so simple that no biochemical complexities need arise. However, for his summer phage course at Cold Spring Harbor, Delbrück required more *mathematics* than most biologists commanded, a requirement enforced by admission tests (Kay, 1993, p. 246). Delbrück became such a powerful organizing force that he was regarded as a kind of godfather and his loyal, brilliant followers a kind of mafia. On the other hand, biochemists initially reacted with suspicion to these physicist interlopers, who were in effect practicing biochemistry "without a license." And the structuralists in England did not follow the information theoretic strategy of Delbrück (Kendrew, 1967).

Delbrück and his cult never did find a fundamental paradox at the root of the duplication problem. Even the Watson-Crick double helix and its implications were disappointing for him, gene duplication being largely a matter of biochemistry: the breaking and reforming of hydrogen bonds. Sensing that the most creative days were nearing an end, Delbrück, already in the early 1950s, was making his team read papers on neurons and the neural system, which he saw as the next big challenge. And, indeed, Stent and others did turn to neurology. For such iconoclasts, molecular genetics was soon to turn into a rather boring academic discipline (Stent, 1968).

The molecular biology story illustrates another point that roughly parallels the business examples. Competition in the sciences is not confined to alternative theories or research programs or even experimental techniques (and their practitioners) within a single field. The more ambitious students and the more creative senior people tend to move to the "hot" areas or to those perceived to ready for creative explosion. As their promise and the challenge of realizing that promise waxes and wanes (together with appropriate funding sources), specialty areas jockey for position. Fledging developments can become front-burner. So, occasionally, can older, neglected problem domains. And vice versa: once hot fields can become backwaters, largely exhausted of their perceived potential for first-rate work. There are many historical examples.

7. CONCLUDING REFLECTIONS: THE INVISIBILITY OF DISRUPTIONS

Disruptions of practices tend to be invisible for the same reasons that Kuhnian revolutions are invisible, plus others. As long as we focus attention on major revolts against established theory, including representations of long-term research programs, we fail to see that theory disruptions are not always in phase with disruptions of practice. The same can happen if our contact with ongoing science is limited to reading refereed mainline publications with their conventionalized descriptions of procedure. Third, experimental, theoretical, and organizational-institutional practices all tend to originate in localized, esoteric contexts. That makes them hard to see.

On the basis of an extensive survey of American market economy since 1900, Foster and Kaplan (2001) show that the half-life of major companies gets shorter and shorter as the market responds ever more quickly to any profit-making advantage that they might have. Is not a similar phenomenon occurring in the sciences, which are also characterized by strong demands for innovation coupled with increasingly intense competition as the number of scientists and their technological capability increase? I have seen no empirical studies of this question, but we might wonder whether "scientific generations" are getting shorter in the sense that the education of new cohorts of students is changing more rapidly than before. One does sense that the windows of opportunity for being first with a new innovation are smaller than before. The best empirical information we have for the sciences as a whole comes from citation analysis. In the introduction to *Structure*, Kuhn already suggested that shifts in citation patterns would indicate the existence of revolutions, large and small.

Some related questions are these. Does Schumpeter's phrase "creative destruction" apply also to scientific work, despite scientists' appeal to nature for empirical evidence? After all, successful businesses as well as successful scientific research programs must produce products widely perceived to work better than those of the competition. Both domains are in some sense progressive even if not in the sense of cumulativity endorsed by the logical empiricists. But in precisely what sense? Was Kuhn basically correct that occasional scientific revolutions, large and small, are virtually inevitable?[25]

I have argued that philosophers need to try harder to escape their own professional mindsets when providing accounts of scientific research. This will require more attention to actual scientific practices in their pragmatic opportunism, with more attention to heuristic appraisal and less to traditional confirmation theory. Although Kuhn took a big step in this direction, he did not go far enough. I have suggested some alternative models, albeit in a very preliminary way.

I conclude by bringing home to general philosophy of science the point about new-market and low-end disruption. From the standpoint of professional philosophy of science and its own continuity assumption, the slowness of philosophers to take seriously actual scientific practices, has allowed science studies to become a disruptive academic technology! Science studies (*sans* philosophy) has captured much of the intellectual market formerly dominated by philosophy of science. The old discovery-justification distinction placed scientific practices out of bounds for traditional philosophy

[25] Inspired by Bak (1996), Mark Buchanan (2001, Chaps. 10, 11), in a popular book on complexity theory and self-organized criticality, intriguing suggests that the kinds of forces are at work in scientific research to predict the existence of scientific revolutions. As with earthquakes and many other phenomena, revolutions large and small should fit a scale-free power distribution law. Buchanan is rightly struck by the remarkable nonlinearity of scientific revolutions: large events may be triggered by small, ordinary-seeming causes.

of science, so there could be no official quarrel with history of science and the new sociology of scientific knowledge (SSK) attempting to fill that void. Laudan (1977), for example, could still voice the position that philosophy of science deals with what is rational in science and leaves the irrational elements to be explained by sociology – at the low end, in other words. Kuhn's use of history to discredit both positivist and Popperian philosophies of science had already caused a great stir when the advocates of SSK issued their manifesto, the Strong Program, that challenged philosophy of science pretty much across the board. Science studies, in its many guises, has expanded rapidly since then. The irony is that context of discovery (understood broadly as that context or set of contexts in which scientists actually do their work at the frontier) is more important to understanding science than context of justification! Of course, science studies practitioners themselves reject the terminology of this distinction.

BIBLIOGRAPHY

Arabatzis, T. (2006). *Representing Electrons: A Biographical Approach to Theoretical Entities*. Chicago, IL: University of Chicago Press.

Bachelard, G. (1934). *Le Nouvel Esprit Scientifique*. Translated as *The New Scientific Spirit*. Boston: Beacon, 1984.

Bak, P. (1996). *How Nature Works*. Oxford: Oxford University Press.

Barker, P., Chen, X., and Andersen, H. (2003). Kuhn on Concepts and Categorization. In T. Nickles, pp. 212–245.

Bloom, H. (1973). *The Anxiety of Influence: A Theory of Poetry*. Oxford: Oxford University Press.

Bohr, N. (1933). Light and Life. *Nature,* 131, 421. Reprinted in Bohr's *Atomic Physics and Human Knowledge*. New York: Science Editions, 1958, pp. 3–12.

Bridgman, P. W. (1927). *The Logic of Modern Physics*. New York: Macmillan.

Buchanan, M. (2001). *Ubiquity: Why Catastrophes Happen*. New York: Three Rivers.

Butterfield, H. (1931). *The Whig Interpretation of History*. London: Macmillan.

Butterfield, H. (1949). *The Origins of Modern Science, 1300–1800*. London: G. Bell. Revised edition, 1957.

Cairns, J., Stent, G. and Watson, J. (eds.) (1992). *Phage and the Origins of Molecular Biology*, expanded edition. Plainview, NY: Cold Spring Harbor Laboratory.

Cartwright, N. (1987). Philosophical Problems of Quantum Theory: The Response of American Physicists. In L. Krüger, G. Gigerenzer, and M. Morgan (eds.) *The Probabilistic Revolution*, Vol. 2. Cambridge, MA: MIT, pp. 417–437.

Christensen, C. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Cambridge, MA: Harvard Business School.

Christensen, C. and Raynor, M. (2003). *The Innovator's Solution*. Cambridge, MA: Harvard Business School.

Davidson, D. (1986). A Nice Derangement of Epitaphs. In Richard Grandy and Richard Warner (eds.) *Philosophical Grounds of Rationality*. Oxford: Oxford University Press, pp. 157–174. Reprinted in Davidson's *Truth, Language and History* (Philosophical Papers, Vol. 5). Oxford: Oxford University Press, 2005.

Delbrück, M. (1949). A Physicist Looks at Biology. As reprinted in Cairns et al. (1992), pp. 9–92.

Dijksterhuis, E. J. (1961). *The Mechanization of the World Picture*. Oxford: Oxford University Press. Originally published in 1950, in Dutch.

Forman, P. (1971). Weimar Culture, Causality, and Quantum Theory, 1918–1927: Adaptation by German Physicists and Mathematicians to a Hostile Intellectual Environment. *Historical Studies in the Physical Sciences,* 2, 10–115.

Foster, R. (1986). *Innovation: The Attacker's Advantage*. New York: Summit/Simon & Schuster.

Foster, R. and Kaplan, S. (2001). *Creative Destruction*. New York: Currency/Doubleday.

Frisch, M. (2005). *Inconsistency, Asymmetry, and Non-Locality: A Philosophical Investigation of Classical Electrodynamics*. Oxford: Oxford University Press.
Galison, P. (1997). *Image and Logic: A Material Culture of Microphysics*. Chicago, IL: University of Chicago Press.
Gigerenzer, G. (2000). *Adaptive Thinking: Rationality in the Real World*. Oxford: Oxford University Press.
Hacking, I. (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.
Hesse, M. (1963). *Models and Analogies in Science*, 2nd ed., London: Sheed & Ward; enlarged, South Bend: University of Notre Dame Press, 1966.
Hook, E. (ed.) (2001). *Prematurity in Scientific Discovery: On Resistance and Neglect*. Berkeley, CA: University of California Press.
Hoyningen-Huene, P. (2006). Context of Discovery versus Context of Justification and Thomas Kuhn. In J. Schickore and F. Steinle (eds.), pp. 119–131.
Hoyningen-Huene, P. and Sankey, H. (eds.) (2001). *Incommensurability and Related Matters*. Dordrecht: Kluwer.
Hull, D. (1989). *The Metaphysics of Evolution*. Albany: SUNY.
James, W. (1880). Great Men, Great Thoughts, and the Environment. *Atlantic Monthly,* 46, 441–459. Reprinted as Great Men and their Environments. In *The Will to Believe and Other Essays*. New York: Dover, pp. 216–254.
James, W. (1896). The Will to Believe. Reprinted in *The Will to Believe and Other Essays in Popular Philosophy*. New York: Dover, 1956.
Kay, L. (1993). *The Molecular Vision of Life: Caltech, the Rockefeller Foundation, and the Rise of the New Biology*. Oxford: Oxford University Press.
Kellert, S. (1993). A Philosophical Evaluation of the Chaos Theory 'Revolution'. In *PSA 1992*, Vol. 2. East Lansing, MI: Philosophy of Science Association, pp. 33–49.
Kendrew, J. (1967). How Molecular Biology Started. *Scientific American,* 216(3), 141–144. As reprinted in Cairns et al. (1992), pp. 343–344.
Kingsland, S. (1995). *Modeling Nature*, 2nd ed. Chicago, IL: University of Chicago Press.
Koyré, A. (1939). *Etudes Galiléennes*. Paris: Hermann. Translated as *Galilean Studies*. Atlantic Highlands, NJ: Humanities Press, 1978.
Kuhn, T. (1955). Carnot's Version of Carnot's Cycle. *American Journal of Physics,* 23, 91–95.
Kuhn, T. (1957). *The Copernican Revolution*. Cambridge, MA: Harvard University Press.
Kuhn, T. (1958). The Caloric Theory of Adiabatic Compression. *Isis,* 49, 132–140.
Kuhn, T. (1961). Sadi Carnot and the Cagnard Engine. *Isis,* 52, 567–574.
Kuhn, T. (1962). *The Structure of Scientific Revolutions*, 2nd ed. Chicago, IL: University of Chicago Press. With Postscript, 1970.
Kuhn, T. (1977). *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago, IL: University of Chicago Press.
Kuhn, T. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894–1912*. Oxford: Clarendon.
Kuhn, T. (2000). A Discussion with Thomas S. Kuhn (with Aristides Baltas, Kostas Gavroglu, and Vassiliki Kindi). In T. Kuhn, pp. 255–323.
Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In Lakatos and Alan Musgrave (eds.) *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, pp. 91–195.
Laudan, L. (1977). *Progress and Its Problems*. Berkeley, CA: University of California Press.
Laudan, L. (1981). A Confutation of Convergent Realism. *Philosophy of Science,* 48, 1–49.
Laudan, L. (1984). *Science and Value*. Berkeley, CA: University of California Press.
Levins, R. (1966). The Strategy of Model Building in Population Biology. *American Scientist*, 54, 421–431. As reprinted in Elliot Sober (ed.) *Conceptual Issues in Evolutionary Biology*. Cambridge, MA: MIT, 1984, pp. 18–27.
Nickles, T. (1992). Good Science Is Bad History. In Ernan McMullin (ed.) *The Social Dimensions of Science*. Notre Dame, IN: University of Notre Dame Press, pp. 895–129.
Nickles, T. (2002). The Discovery-Justification (D-J) Distinction and Professional Philosophy of Science: Comments on the First Day's Five Papers. In Jutta Schickore and Friedrich Steinle (eds.) *Revisiting*

Discovery and Justification. Preprint 211. Berlin: Max-Planck-Institut für Wissenschaftsgeschichte, pp. 67–78.

Nickles, T. (2003a). Normal Science: From Logic to Case-Based and Model-Based Reasoning. In T. Nickles (ed.), pp. 142–177.

Nickles, T, (ed.) (2003b.) *Thomas Kuhn*. Cambridge: Cambridge University Press.

Nickles, T. (2006a.) Scientific Revolutions. In Sahotra S. and Pfeifer J. (eds.) *The Philosophy of Science: An Encyclopedia*, Vol. 2. New York: Routledge, pp. 754–765.

Nickles, T. (2006b). Heuristic Appraisal: Context of Discovery or Justification? In J. Schickore and Friedrich S. (eds.) *Revisiting Discovery and Justification: Historical and Philosophical Perspectives on the Context Distinction*. Dordrecht, The Netherlands: Springer, pp. 159–182.

Pickering, A. (1980a). Exemplars and Analogies: A Comment on Crane's Study of Kuhnian Paradigms in High Energy Physics and Reply to Crane. *Social Studies of Science,* 10: 497–502 and 507–508.

Pickering, A. (1980b). The Role of Interests in High-Energy Physics: The Choice between Charm and Colour. In K. Knorr, R. Krohn, and R. Whitley (eds.), *Sociology of the Sciences Yearbook 1980*. Dordrecht: Reidel, pp. 107–138.

Pickering, A. (1984). *Constructing Quarks: A Sociological History of Particle Physics*. Chicago, IL: University of Chicago Press.

Poincaré, H. (1963). *Mathematics and Science: Last Essays*. New York: Dover. Translated from the French edition of 1913.

Price, D. (1963). *Little Science, Big Science*. New York: Columbia University Press.

Rheinberger, H.-J. (1997). *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*. Stanford: Stanford University Press.

Rohrlich, F. (1965). *Classical Charged Particles*. Reading, MA: Addison-Wesley.

Rouse, J. (1996). *Engaging Science: How to Understand Its Practices Philosophically*. Ithaca: Cornell University Press.

Rouse, J. (2003). Kuhn's Philosophy of Scientific Practice. In T. Nickles (ed.) pp. 101–121.

Ruse, M. (2001). The Prematurity of Darwin's Theory of Natural Selection. In E. Hook (ed.) pp. 215–238.

Schilpp, P. A. (ed.) (1949). *Albert Einstein: Philosopher-Scientist*. Evanston: Library of Living Philosophers.

Schrödinger, E. (1945). *What Is Life?* Cambridge: Cambridge University Press.

Schumpeter, J. (1942). *Capitalism, Socialism, and Democracy*. New York: Harper.

Shapere, D. (1984). *Reason and the Search for Knowledge*. Dordrecht, The Netherlands: Reidel, pp. 383–407.

Shapere, D. (2001). Reasons, Radical Change and Incommensurability in Science. In P. Hoyningen and H. Sankey (ed.) pp. 181–206.

Stent, G. (1968). That Was the Molecular Biology That Was. *Science,* 160: 390–395. Reprinted in Cairns et al. (1992) pp. 345–350.

Stent, G. (1971). *Molecular Genetics: An Introductory Narrative*. San Francisco: W. H. Freeman.

Toulmin, S. (1972). *Human Understanding*. Princeton, NJ: Princeton University Press.

Toulmin, S. (1988). Introduction. *John Dewey, Vol. 4, 1929: The Quest for Certainty*. In Jo Ann Boydston (ed.) Carbondale, IL: Southern Illinois University Press, pp. vii–xxii.

Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer. English translation, Princeton University Press, 1955.

Wimsatt, W. C. (1987). False Models as Means to Truer Theories. In M. Nitecki and A. Hoffman (eds.), *Neutral Models in Biology*. Oxford: Oxford University Press, pp. 23–55.

Zuckerman, H. and J. Lederberg (1986). Postmature Scientific Discovery. *Nature,* 324, 629–631.

SCIENTIFIC REVOLUTIONS: THE VIEW FROM INSIDE AND THE VIEW FROM OUTSIDE

Commentary on "Disruptive Scientific Change", by Thomas Nickles

EMILIANO TRIZIO

Thomas Nickles' central thesis, namely that "there is a double asymmetry in the importance of incommensurability as it concerns philosophers and scientists" (p. 352) invites us to reflect on the status of the philosophical and historical accounts of science. Traditional philosophers of science, along with philosophically minded historians, are criticized on the grounds that they surreptitiously identify scientists' problems and concerns with their own. According to Thomas Nickles, Kuhn himself, who has contributed probably more than anyone else to the rise of historical philosophy of science, is not exempt from this criticism: his effort to cast light on the true nature of science, by looking at historical records and by reconstructing the *invisible* revolutionary structure they seem to reveal, has been partly vitiated, if I understand correctly, by two kinds of "professional biases." Given that Kuhn and his followers were philosophically minded historians, it is not surprising that their biases should derive partly from their philosophical interests and partly from the requirements, and even the constraints, that characterize historical researches as such. As far as philosophy is concerned, the so-called historical turn has been mainly motivated by the intent to overthrow the "received view" about science developed by the logical empiricists. Scientific revolutions and incommensurability have thus become the watchwords of those who aimed at showing that the *Wissenschaftslogik* was unable to give a satisfactory account of scientific knowledge. Indeed the shift from the lofty realm of formal logic and semantics to the more down-to-earth vicissitudes of the history of science was still motivated by philosophical worries exogenous to the actual preoccupations of working scientists. The main consequence of this dialectical faithfulness to the agenda set by the logical empiricists has been an undue insistence on the representational character of both continuity and discontinuity in the evolution of science. Let us now turn to the second set of biases, namely those deriving from the historical research. As Nickles says, historians write narratives and the very structure of narratives encourages them to emphasize whatever allows the arrangement of huge amounts of little events and details in a coherent, well-structured and readable form. Sharp breaks, rapid upheavals followed by Orwellian occultation of the loser's good reasons provide the narrative frameworks needed by historians. Now, Nickles' central thesis can be summarized by saying that scientists do not really *experience* these kinds of representational breakdowns,

L. Soler, H. Sankey and P. Hoyningen-Huene, (eds.), Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities, 381–384.

they adapt rather easily to representational change and disagreement at the verbal level while, instead, they find it difficult to cope with "disruptive change" at the level of the practices characterizing a certain research field. This is, in turn, a consequence of the fact that researchers are less interested in the representational content of science, than in the heuristic and potential dimension of the available ideas, procedures and technical means, which have to constitute viable and promising *tools* in the open and project-laden horizon of research. We could say that for a today's researcher scientific work is successful mainly insofar it leads to *more* scientific work. A double shift of the focus of philosophical analysis is thus needed, on the one hand from the context of justification to the context of discovery and, on the other hand, from representational contents to practices. I will now try to make a few related remarks.

Do scientists experience scientific revolutions? I will, first of all, focus on the problem of the discrepancy between the historical and philosophical accounts of scientific revolutions and the actual experiences of working scientists. As a matter of fact, Kuhn has gradually come to take into account this problem. I will quote *passim* a passage taken from *Commensurability, Comparability, Communicability*: "The concept of scientific revolution originated in the discovery that to understand any part of the science of the past the historian must first learn the language in which the past was written. […] Since success in interpretation is generally achieved in large chunks […], the historian's discovery of the past repeatedly involves the sudden recognition of new patterns or gestalts. It follows that the historian, at least, does experience revolutions. […] Whether scientists, moving through time in a direction opposite to the historian's, also experience revolutions is left open by what I have so far said. If they do, their shifts in gestalt will ordinarily be smaller than the historian's, for what the latter experiences as a single revolutionary change will usually have been spread over a number of such changes during the development of the sciences."[1] Kuhn then considers the possibility that what actually takes place in history are gradual linguistic drifts, but he plays it down on the grounds of the historical evidence and on the grounds of his general views about scientific language. He nevertheless concludes that "If I were now rewriting *The Structure of Scientific Revolutions*, I would emphasize language change more and the normal/revolutionary distinction less."[2] To be entirely fair to Kuhn, we must therefore acknowledge that he was not blind to the fact that the historian's point of view introduces some biases in the description of the development of science. As Thomas Nickles says, it is a complex interpretative problem to understand just to what extent Kuhn has anticipated the more recent interest for scientific practices; nevertheless, it does seem that Kuhn's awareness of the dangers of the historical perspectives on science has not led him to give up his excessive stress on representations and representational breaks. Indeed the attempt to develop a satisfactory account of incommensurability has led Kuhn to treat scientific communities more and more as *language* communities. We can conclude that in Kuhn's work the philosophical biases inherited from the logical empiricists have been, in the long run, more harmful and more unconsciously active than those determined by the historical research itself.

1 Kuhn (2000, pp. 56–57).
2 Ibid.

Unintelligibility. In order to lay the foundations of an account of scientific disruptions more faithful to the real scientific practices, Thomas Nickles proposes an interesting definition of "unintelligibility." He writes, "there *is* something that could be called unintelligibility within scientific work besides an inability to make sense of a claim or a practice. It is an inability of scientific specialists to see the *relevance* of a research claim, technique, or strategy to their field, to their expertise, to their own specific work. As scientists, they cannot figure out anything important that they can *do* with it. […] In short, these specialists fail to *recognize* it as a contribution to their field" (p. 353). Thus construed, unintelligibility is a recognition problem. Indeed it seems to me that there are two different possibilities that are suggested by the preceding quotation and that should be distinguished. In virtue of its intellectual understanding of a novelty, a scientist may well grant that it belongs to his or her field of investigation and yet deny that it is *helpful* in any sense (for practical or theoretical reasons). On the other hand, a novelty may be considered outright *irrelevant* for a given discipline, simply not belonging to it. From the point of view of the working scientist, the two situations correspond respectively to the question about the *fruitfulness* and the question about the *relevance* of a novelty. Both situations have to do with the way researchers cope with new proposals and may affect in the long run the integrity of a discipline, although only the second one explicitly implies a reconsideration of its boundaries. It is, in general, the consideration of this kind of heuristic factors that constitutes a powerful tool, within Nickles' scheme, for analysing the disruptive potential of new procedures and techniques that do not necessarily emerge in times of deep conceptual revolutions. This paves the way, in turn, to an account of the development of new specialties and subspecialties that is more fluid and, probably, more realistic than the one proposed by Kuhn in his paper *The Road Since Structure*, which is based on the specialisation of the scientific lexicon that occurs after a scientific (and hence, conceptual) revolution.[3]

A plea for philosophers. In conclusion, I would like to say that I do not interpret Thomas Nickles' contribution to the understanding of scientific practices as necessarily implying a criticism of philosophy of science as such. The philosophers' biases are harmful only insofar as they lead to wrong interpretations of scientific practices and tend to attribute to the researchers the problems that derive from their own concerns. Nevertheless, there is more to philosophy of science than the task of helping historians and sociologists to provide a faithful account of what real science is and how scientists really work. What I take to be its most important issue, namely understanding what kind of grasp science has on reality, may demand that a dialogue between Aristotle, Newton and Einstein be imagined, although it would not help us in any way to understand how scientists experience disruptive change. Moreover, the shift from a theory-centred to a practice-centred philosophical account of science, far from eliminating the problem of scientific realism, allows us to reconsider it in a more suitable theoretical framework. In short, we are interested both in a *view from within* and in a *view from outside* science: Thomas Nickles has warned us not to conflate the two analytical levels, as Kuhn, sometimes, probably did.

[3] Ibid., pp. 97–99.

BIBLIOGRAPHY

Kuhn, T. (1983) Commensurability, Comparability, Communicability. In P. D. Asquith and T. Nickles (eds.) *PSA 1982: Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association*, Vol. 2, East Lansing, MI: Philosophy of Science association, pp. 669–688. Reprinted in T. Kuhn (2000). *The Road since Structure*, University of Chicago Press, Chicago, IL, pp. 33–57.

Kuhn, T. (1991) The Road since Structure. In A. Fine, M. Forbes, and L. Wessels (eds.) *PSA 1990: Proceedings of the 1990 Biennial Meeting of the Philosophy of Science Association*, Vol. 2, East Lansing, MI: Philosophy of Science association, pp. 3–13. Reprinted in T. Kuhn (2000). *The Road since Structure*, University of Chicago Press, Chicago, IL, pp. 90–104.

Kuhn, T. (2000) *The Road since Structure*. Chicago, IL: University of Chicago Press.

Boston Studies in the Philosophy of Science

Editor: Robert S. Cohen, *Boston University*

1. M.W. Wartofsky (ed.): *Proceedings of the Boston Colloquium for the Philosophy of Science, 1961/1962.* [Synthese Library 6] 1963 ISBN 90-277-0021-4
2. R.S. Cohen and M.W. Wartofsky (eds.): *Proceedings of the Boston Colloquium for the Philosophy of Science, 1962/1964.* In Honor of P. Frank. [Synthese Library 10] 1965 ISBN 90-277-9004-0
3. R.S. Cohen and M.W. Wartofsky (eds.): *Proceedings of the Boston Colloquium for the Philosophy of Science, 1964/1966.* In Memory of Norwood Russell Hanson. [Synthese Library 14] 1967 ISBN 90-277-0013-3
4. R.S. Cohen and M.W. Wartofsky (eds.): *Proceedings of the Boston Colloquium for the Philosophy of Science, 1966/1968.* [Synthese Library 18] 1969 ISBN 90-277-0014-1
5. R.S. Cohen and M.W. Wartofsky (eds.): *Proceedings of the Boston Colloquium for the Philosophy of Science, 1966/1968.* [Synthese Library 19] 1969 ISBN 90-277-0015-X
6. R.S. Cohen and R.J. Seeger (eds.): *Ernst Mach, Physicist and Philosopher.* [Synthese Library 27] 1970 ISBN 90-277-0016-8
7. M. Čapek: *Bergson and Modern Physics.* A Reinterpretation and Re-evaluation. [Synthese Library 37] 1971 ISBN 90-277-0186-5
8. R.C. Buck and R.S. Cohen (eds.): *PSA 1970.* Proceedings of the 2nd Biennial Meeting of the Philosophy and Science Association (Boston, Fall 1970). In Memory of Rudolf Carnap. [Synthese Library 39] 1971 ISBN 90-277-0187-3; Pb 90-277-0309-4
9. A.A. Zinov'ev: *Foundations of the Logical Theory of Scientific Knowledge (Complex Logic).* Translated from Russian. Revised and enlarged English Edition, with an Appendix by G.A. Smirnov, E.A. Sidorenko, A.M. Fedina and L.A. Bobrova. [Synthese Library 46] 1973 ISBN 90-277-0193-8; Pb 90-277-0324-8
10. L. Tondl: *Scientific Procedures.* A Contribution Concerning the Methodological Problems of Scientific Concepts and Scientific Explanation. Translated from Czech. [Synthese Library 47] 1973 ISBN 90-277-0147-4; Pb 90-277-0323-X
11. R.J. Seeger and R.S. Cohen (eds.): *Philosophical Foundations of Science.* Proceedings of Section L, 1969, American Association for the Advancement of Science. [Synthese Library 58] 1974 ISBN 90-277-0390-6; Pb 90-277-0376-0
12. A. Grünbaum: *Philosophical Problems of Space and Times.* 2nd enlarged ed. [Synthese Library 55] 1973 ISBN 90-277-0357-4; Pb 90-277-0358-2
13. R.S. Cohen and M.W. Wartofsky (eds.): *Logical and Epistemological Studies in Contemporary Physics.* Proceedings of the Boston Colloquium for the Philosophy of Science, 1969/72, Part I. [Synthese Library 59] 1974 ISBN 90-277-0391-4; Pb 90-277-0377-9
14. R.S. Cohen and M.W. Wartofsky (eds.): *Methodological and Historical Essays in the Natural and Social Sciences.* Proceedings of the Boston Colloquium for the Philosophy of Science, 1969/72, Part II. [Synthese Library 60] 1974 ISBN 90-277-0392-2; Pb 90-277-0378-7
15. R.S. Cohen, J.J. Stachel and M.W. Wartofsky (eds.): *For Dirk Struik.* Scientific, Historical and Political Essays in Honor of Dirk J. Struik. [Synthese Library 61] 1974 ISBN 90-277-0393-0; Pb 90-277-0379-5
16. N. Geschwind: *Selected Papers on Language and the Brains.* [Synthese Library 68] 1974 ISBN 90-277-0262-4; Pb 90-277-0263-2
17. B.G. Kuznetsov: *Reason and Being.* Translated from Russian. Edited by C.R. Fawcett and R.S. Cohen. 1987 ISBN 90-277-2181-5
18. P. Mittelstaedt: *Philosophical Problems of Modern Physics.* Translated from the revised 4th German edition by W. Riemer and edited by R.S. Cohen. [Synthese Library 95] 1976 ISBN 90-277-0285-3; Pb 90-277-0506-2

19. H. Mehlberg: *Time, Causality, and the Quantum Theory.* Studies in the Philosophy of Science. Vol. I: *Essay on the Causal Theory of Time.* Vol. II: *Time in a Quantized Universe.* Translated from French. Edited by R.S. Cohen. 1980 Vol. I: ISBN 90-277-0721-9; Pb 90-277-1074-0
Vol. II: ISBN 90-277-1075-9; Pb 90-277-1076-7
20. K.F. Schaffner and R.S. Cohen (eds.): *PSA 1972.* Proceedings of the 3rd Biennial Meeting of the Philosophy of Science Association (Lansing, Michigan, Fall 1972). [Synthese Library 64] 1974 ISBN 90-277-0408-2; Pb 90-277-0409-0
21. R.S. Cohen and J.J. Stachel (eds.): *Selected Papers of Léon Rosenfeld.* [Synthese Library 100] 1979 ISBN 90-277-0651-4; Pb 90-277-0652-2
22. M. Čapek (ed.): *The Concepts of Space and Time.* Their Structure and Their Development. [Synthese Library 74] 1976 ISBN 90-277-0355-8; Pb 90-277-0375-2
23. M. Grene: *The Understanding of Nature.* Essays in the Philosophy of Biology. [Synthese Library 66] 1974 ISBN 90-277-0462-7; Pb 90-277-0463-5
24. D. Ihde: *Technics and Praxis.* A Philosophy of Technology. [Synthese Library 130] 1979 ISBN 90-277-0953-X; Pb 90-277-0954-8
25. J. Hintikka and U. Remes: *The Method of Analysis.* Its Geometrical Origin and Its General Significance. [Synthese Library 75] 1974 ISBN 90-277-0532-1; Pb 90-277-0543-7
26. J.E. Murdoch and E.D. Sylla (eds.): *The Cultural Context of Medieval Learning.* Proceedings of the First International Colloquium on Philosophy, Science, and Theology in the Middle Ages, 1973. [Synthese Library 76] 1975 ISBN 90-277-0560-7; Pb 90-277-0587-9
27. M. Grene and E. Mendelsohn (eds.): *Topics in the Philosophy of Biology.* [Synthese Library 84] 1976 ISBN 90-277-0595-X; Pb 90-277-0596-8
28. J. Agassi: *Science in Flux.* [Synthese Library 80] 1975 ISBN 90-277-0584-4; Pb 90-277-0612-3
29. J.J. Wiatr (ed.): *Polish Essays in the Methodology of the Social Sciences.* [Synthese Library 131] 1979 ISBN 90-277-0723-5; Pb 90-277-0956-4
30. P. Janich: *Protophysics of Time.* Constructive Foundation and History of Time Measurement. Translated from German. 1985 ISBN 90-277-0724-3
31. R.S. Cohen and M.W. Wartofsky (eds.): *Language, Logic, and Method.* 1983 ISBN 90-277-0725-1
32. R.S. Cohen, C.A. Hooker, A.C. Michalos and J.W. van Evra (eds.): *PSA 1974.* Proceedings of the 4th Biennial Meeting of the Philosophy of Science Association. [Synthese Library 101] 1976 ISBN 90-277-0647-6; Pb 90-277-0648-4
33. G. Holton and W.A. Blanpied (eds.): *Science and Its Public.* The Changing Relationship. [Synthese Library 96] 1976 ISBN 90-277-0657-3; Pb 90-277-0658-1
34. M.D. Grmek, R.S. Cohen and G. Cimino (eds.): *On Scientific Discovery.* The 1977 Erice Lectures. 1981 ISBN 90-277-1122-4; Pb 90-277-1123-2
35. S. Amsterdamski: *Between Experience and Metaphysics.* Philosophical Problems of the Evolution of Science. Translated from Polish. [Synthese Library 77] 1975 ISBN 90-277-0568-2; Pb 90-277-0580-1
36. M. Marković and G. Petrović (eds.): *Praxis.* Yugoslav Essays in the Philosophy and Methodology of the Social Sciences. [Synthese Library 134] 1979 ISBN 90-277-0727-8; Pb 90-277-0968-8
37. H. von Helmholtz: *Epistemological Writings.* The Paul Hertz / Moritz Schlick Centenary Edition of 1921. Translated from German by M.F. Lowe. Edited with an Introduction and Bibliography by R.S. Cohen and Y. Elkana. [Synthese Library 79] 1977 ISBN 90-277-0290-X; Pb 90-277-0582-8
38. R.M. Martin: *Pragmatics, Truth and Language.* 1979 ISBN 90-277-0992-0; Pb 90-277-0993-9
39. R.S. Cohen, P.K. Feyerabend and M.W. Wartofsky (eds.): *Essays in Memory of Imre Lakatos.* [Synthese Library 99] 1976 ISBN 90-277-0654-9; Pb 90-277-0655-7
40. Not published.

41. Not published.
42. H.R. Maturana and F.J. Varela: *Autopoiesis and Cognition.* The Realization of the Living.With a Preface to “Autopoiesis’ by S. Beer. 1980 ISBN 90-277-1015-5; Pb 90-277-1016-3
43. A. Kasher (ed.): *Language in Focus: Foundations, Methods and Systems.* Essays in Memory of Yehoshua Bar-Hillel. [Synthese Library 89] 1976 ISBN 90-277-0644-1; Pb 90-277-0645-X
44. T.D. Thao: *Investigations into the Origin of Language and Consciousness.* 1984 ISBN 90-277-0827-4
45. F.G.-I. Nagasaka (ed.): *Japanese Studies in the Philosophy of Science.* 1997 ISBN 0-7923-4781-1
46. P.L. Kapitza: *Experiment, Theory, Practice.* Articles and Addresses. Edited by R.S. Cohen. 1980 ISBN 90-277-1061-9; Pb 90-277-1062-7
47. M.L. Dalla Chiara (ed.): *Italian Studies in the Philosophy of Science.* 1981 ISBN 90-277-0735-9; Pb 90-277-1073-2
48. M.W. Wartofsky: *Models.* Representation and the Scientific Understanding. [Synthese Library 129] 1979 ISBN 90-277-0736-7; Pb 90-277-0947-5
49. T.D. Thao: *Phenomenology and Dialectical Materialism.* Edited by R.S. Cohen. 1986 ISBN 90-277-0737-5
50. Y. Fried and J. Agassi: *Paranoia.* A Study in Diagnosis. [Synthese Library 102] 1976 ISBN 90-277-0704-9; Pb 90-277-0705-7
51. K.H. Wolff: *Surrender and Cath.* Experience and Inquiry Today. [Synthese Library 105] 1976 ISBN 90-277-0758-8; Pb 90-277-0765-0
52. K. Kosík: *Dialectics of the Concrete.* A Study on Problems of Man and World. 1976 ISBN 90-277-0761-8; Pb 90-277-0764-2
53. N. Goodman: *The Structure of Appearance.* [Synthese Library 107] 1977 ISBN 90-277-0773-1; Pb 90-277-0774-X
54. H.A. Simon: *Models of Discovery* and Other Topics in the Methods of Science. [Synthese Library 114] 1977 ISBN 90-277-0812-6; Pb 90-277-0858-4
55. M. Lazerowitz: *The Language of Philosophy.* Freud and Wittgenstein. [Synthese Library 117] 1977 ISBN 90-277-0826-6; Pb 90-277-0862-2
56. T. Nickles (ed.): *Scientific Discovery, Logic, and Rationality.* 1980 ISBN 90-277-1069-4; Pb 90-277-1070-8
57. J. Margolis: *Persons and Mind.* The Prospects of Nonreductive Materialism. [Synthese Library 121] 1978 ISBN 90-277-0854-1; Pb 90-277-0863-0
58. G. Radnitzky and G. Andersson (eds.): *Progress and Rationality in Science.* [Synthese Library 125] 1978 ISBN 90-277-0921-1; Pb 90-277-0922-X
59. G. Radnitzky and G. Andersson (eds.): *The Structure and Development of Science.* [Synthese Library 136] 1979 ISBN 90-277-0994-7; Pb 90-277-0995-5
60. T. Nickles (ed.): *Scientific Discovery.* Case Studies. 1980 ISBN 90-277-1092-9; Pb 90-277-1093-7
61. M.A. Finocchiaro: *Galileo and the Art of Reasoning.* Rhetorical Foundation of Logic and Scientific Method. 1980 ISBN 90-277-1094-5; Pb 90-277-1095-3
62. W.A. Wallace: *Prelude to Galileo.* Essays on Medieval and 16th-Century Sources of Galileo’s Thought. 1981 ISBN 90-277-1215-8; Pb 90-277-1216-6
63. F. Rapp: *Analytical Philosophy of Technology.* Translated from German. 1981 ISBN 90-277-1221-2; Pb 90-277-1222-0
64. R.S. Cohen and M.W. Wartofsky (eds.): *Hegel and the Sciences.* 1984 ISBN 90-277-0726-X
65. J. Agassi: *Science and Society.* Studies in the Sociology of Science. 1981 ISBN 90-277-1244-1; Pb 90-277-1245-X
66. L. Tondl: *Problems of Semantics.* A Contribution to the Analysis of the Language of Science. Translated from Czech. 1981 ISBN 90-277-0148-2; Pb 90-277-0316-7

Boston Studies in the Philosophy of Science

67. J. Agassi and R.S. Cohen (eds.): *Scientific Philosophy Today.* Essays in Honor of Mario Bunge. 1982 ISBN 90-277-1262-X; Pb 90-277-1263-8
68. W. Krajewski (ed.): *Polish Essays in the Philosophy of the Natural Sciences.* Translated from Polish and edited by R.S. Cohen and C.R. Fawcett. 1982 ISBN 90-277-1286-7; Pb 90-277-1287-5
69. J.H. Fetzer: *Scientific Knowledge.* Causation, Explanation and Corroboration. 1981 ISBN 90-277-1335-9; Pb 90-277-1336-7
70. S. Grossberg: *Studies of Mind and Brain.* Neural Principles of Learning, Perception, Development, Cognition, and Motor Control. 1982 ISBN 90-277-1359-6; Pb 90-277-1360-X
71. R.S. Cohen and M.W. Wartofsky (eds.): *Epistemology, Methodology, and the Social Sciences.* 1983 ISBN 90-277-1454-1
72. K. Berka: *Measurement.* Its Concepts, Theories and Problems. Translated from Czech. 1983 ISBN 90-277-1416-9
73. G.L. Pandit: *The Structure and Growth of Scientific Knowledge.* A Study in the Methodology of Epistemic Appraisal. 1983 ISBN 90-277-1434-7
74. A.A. Zinov'ev: *Logical Physics.* Translated from Russian. Edited by R.S. Cohen. 1983 [*see also* Volume 9] ISBN 90-277-0734-0
75. G-G. Granger: *Formal Thought and the Sciences of Man.* Translated from French. With and Introduction by A. Rosenberg. 1983 ISBN 90-277-1524-6
76. R.S. Cohen and L. Laudan (eds.): *Physics, Philosophy and Psychoanalysis.* Essays in Honor of Adolf Grünbaum. 1983 ISBN 90-277-1533-5
77. G. Böhme, W. van den Daele, R. Hohlfeld, W. Krohn and W. Schäfer: *Finalization in Science.* The Social Orientation of Scientific Progress. Translated from German. Edited by W. Schäfer. 1983 ISBN 90-277-1549-1
78. D. Shapere: *Reason and the Search for Knowledge.* Investigations in the Philosophy of Science. 1984 ISBN 90-277-1551-3; Pb 90-277-1641-2
79. G. Andersson (ed.): *Rationality in Science and Politics.* Translated from German. 1984 ISBN 90-277-1575-0; Pb 90-277-1953-5
80. P.T. Durbin and F. Rapp (eds.): *Philosophy and Technology.* [*Also* Philosophy and Technology Series, Vol. 1] 1983 ISBN 90-277-1576-9
81. M. Marković: *Dialectical Theory of Meaning.* Translated from Serbo-Croat. 1984 ISBN 90-277-1596-3
82. R.S. Cohen and M.W. Wartofsky (eds.): *Physical Sciences and History of Physics.* 1984 ISBN 90-277-1615-3
83. É. Meyerson: *The Relativistic Deduction.* Epistemological Implications of the Theory of Relativity. Translated from French. With a Review by Albert Einstein and an Introduction by Milić Čapek. 1985 ISBN 90-277-1699-4
84. R.S. Cohen and M.W. Wartofsky (eds.): *Methodology, Metaphysics and the History of Science.* In Memory of Benjamin Nelson. 1984 ISBN 90-277-1711-7
85. G. Tamás: *The Logic of Categories.* Translated from Hungarian. Edited by R.S. Cohen. 1986 ISBN 90-277-1742-7
86. S.L. de C. Fernandes: *Foundations of Objective Knowledge.* The Relations of Popper's Theory of Knowledge to That of Kant. 1985 ISBN 90-277-1809-1
87. R.S. Cohen and T. Schnelle (eds.): *Cognition and Fact.* Materials on Ludwik Fleck. 1986 ISBN 90-277-1902-0
88. G. Freudenthal: Atom and Individual in the Age of Newton. On the Genesis of the Mechanistic World View. Translated from German. 1986 ISBN 90-277-1905-5
89. A. Donagan, A.N. Perovich Jr and M.V. Wedin (eds.): Human Nature and Natural Knowledge. Essays presented to Marjorie Grene on the Occasion of Her 75th Birthday. 1986 ISBN 90-277-1974-8

90. C. Mitcham and A. Hunning (eds.): *Philosophy and Technology II.* Information Technology and Computers in Theory and Practice. [*Also* Philosophy and Technology Series, Vol. 2] 1986 ISBN 90-277-1975-6
91. M. Grene and D. Nails (eds.): *Spinoza and the Sciences.* 1986 ISBN 90-277-1976-4
92. S.P. Turner: *The Search for a Methodology of Social Science.* Durkheim, Weber, and the 19th-Century Problem of Cause, Probability, and Action. 1986 ISBN 90-277-2067-3
93. I.C. Jarvie: *Thinking about Society.* Theory and Practice. 1986 ISBN 90-277-2068-1
94. E. Ullmann-Margalit (ed.): *The Kaleidoscope of Science.* The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Vol. 1. 1986 ISBN 90-277-2158-0; Pb 90-277-2159-9
95. E. Ullmann-Margalit (ed.): *The Prism of Science.* The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Vol. 2. 1986 ISBN 90-277-2160-2; Pb 90-277-2161-0
96. G. Márkus: *Language and Production.* A Critique of the Paradigms. Translated from French. 1986 ISBN 90-277-2169-6
97. F. Amrine, F.J. Zucker and H. Wheeler (eds.): *Goethe and the Sciences: A Reappraisal.* 1987 ISBN 90-277-2265-X; Pb 90-277-2400-8
98. J.C. Pitt and M. Pera (eds.): *Rational Changes in Science.* Essays on Scientific Reasoning. Translated from Italian. 1987 ISBN 90-277-2417-2
99. O. Costa de Beauregard: *Time, the Physical Magnitude.* 1987 ISBN 90-277-2444-X
100. A. Shimony and D. Nails (eds.): *Naturalistic Epistemology.* A Symposium of Two Decades. 1987 ISBN 90-277-2337-0
101. N. Rotenstreich: *Time and Meaning in History.* 1987 ISBN 90-277-2467-9
102. D.B. Zilberman: *The Birth of Meaning in Hindu Thought.* Edited by R.S. Cohen. 1988 ISBN 90-277-2497-0
103. T.F. Glick (ed.): *The Comparative Reception of Relativity.* 1987 ISBN 90-277-2498-9
104. Z. Harris, M. Gottfried, T. Ryckman, P. Mattick Jr, A. Daladier, T.N. Harris and S. Harris: *The Form of Information in Science.* Analysis of an Immunology Sublanguage. With a Preface by Hilary Putnam. 1989 ISBN 90-277-2516-0
105. F. Burwick (ed.): *Approaches to Organic Form.* Permutations in Science and Culture. 1987 ISBN 90-277-2541-1
106. M. Almási: *The Philosophy of Appearances.* Translated from Hungarian. 1989 ISBN 90-277-2150-5
107. S. Hook, W.L. O'Neill and R. O'Toole (eds.): *Philosophy, History and Social Action.* Essays in Honor of Lewis Feuer. With an Autobiographical Essay by L. Feuer. 1988 ISBN 90-277-2644-2
108. I. Hronszky, M. Fehér and B. Dajka: *Scientific Knowledge Socialized.* Selected Proceedings of the 5th Joint International Conference on the History and Philosophy of Science organized by the IUHPS (Veszprém, Hungary, 1984). 1988 ISBN 90-277-2284-6
109. P. Tillers and E.D. Green (eds.): *Probability and Inference in the Law of Evidence.* The Uses and Limits of Bayesianism. 1988 ISBN 90-277-2689-2
110. E. Ullmann-Margalit (ed.): *Science in Reflection.* The Israel Colloquium: Studies in History, Philosophy, and Sociology of Science, Vol. 3. 1988 ISBN 90-277-2712-0; Pb 90-277-2713-9
111. K. Gavroglu, Y. Goudaroulis and P. Nicolacopoulos (eds.): *Imre Lakatos and Theories of Scientific Change.* 1989 ISBN 90-277-2766-X
112. B. Glassner and J.D. Moreno (eds.): *The Qualitative-Quantitative Distinction in the Social Sciences.* 1989 ISBN 90-277-2829-1
113. K. Arens: *Structures of Knowing.* Psychologies of the 19th Century. 1989 ISBN 0-7923-0009-2
114. A. Janik: *Style, Politics and the Future of Philosophy.* 1989 ISBN 0-7923-0056-4

Boston Studies in the Philosophy of Science

115. F. Amrine (ed.): *Literature and Science as Modes of Expression.* With an Introduction by S. Weininger. 1989 ISBN 0-7923-0133-1
116. J.R. Brown and J. Mittelstrass (eds.): *An Intimate Relation.* Studies in the History and Philosophy of Science. Presented to Robert E. Butts on His 60th Birthday. 1989 ISBN 0-7923-0169-2
117. F. D'Agostino and I.C. Jarvie (eds.): *Freedom and Rationality.* Essays in Honor of John Watkins. 1989 ISBN 0-7923-0264-8
118. D. Zolo: *Reflexive Epistemology.* The Philosophical Legacy of Otto Neurath. 1989 ISBN 0-7923-0320-2
119. M. Kearn, B.S. Philips and R.S. Cohen (eds.): *Georg Simmel and Contemporary Sociology.* 1989 ISBN 0-7923-0407-1
120. T.H. Levere and W.R. Shea (eds.): *Nature, Experiment and the Science.* Essays on Galileo and the Nature of Science. In Honour of Stillman Drake. 1989 ISBN 0-7923-0420-9
121. P. Nicolacopoulos (ed.): *Greek Studies in the Philosophy and History of Science.* 1990 ISBN 0-7923-0717-8
122. R. Cooke and D. Costantini (eds.): *Statistics in Science.* The Foundations of Statistical Methods in Biology, Physics and Economics. 1990 ISBN 0-7923-0797-6
123. P. Duhem: *The Origins of Statics.* Translated from French by G.F. Leneaux, V.N. Vagliente and G.H. Wagner. With an Introduction by S.L. Jaki. 1991 ISBN 0-7923-0898-0
124. H. Kamerlingh Onnes: *Through Measurement to Knowledge.* The Selected Papers, 1853-1926. Edited and with an Introduction by K. Gavroglu and Y. Goudaroulis. 1991 ISBN 0-7923-0825-5
125. M. Čapek: *The New Aspects of Time: Its Continuity and Novelties.* Selected Papers in the Philosophy of Science. 1991 ISBN 0-7923-0911-1
126. S. Unguru (ed.): *Physics, Cosmology and Astronomy, 1300–1700.* Tension and Accommodation. 1991 ISBN 0-7923-1022-5
127. Z. Bechler: *Newton's Physics on the Conceptual Structure of the Scientific Revolution.* 1991 ISBN 0-7923-1054-3
128. É. Meyerson: *Explanation in the Sciences.* Translated from French by M-A. Siple and D.A. Siple. 1991 ISBN 0-7923-1129-9
129. A.I. Tauber (ed.): *Organism and the Origins of Self.* 1991 ISBN 0-7923-1185-X
130. F.J. Varela and J-P. Dupuy (eds.): *Understanding Origins.* Contemporary Views on the Origin of Life, Mind and Society. 1992 ISBN 0-7923-1251-1
131. G.L. Pandit: *Methodological Variance.* Essays in Epistemological Ontology and the Methodology of Science. 1991 ISBN 0-7923-1263-5
132. G. Munévar (ed.): *Beyond Reason.* Essays on the Philosophy of Paul Feyerabend. 1991 ISBN 0-7923-1272-4
133. T.E. Uebel (ed.): *Rediscovering the Forgotten Vienna Circle.* Austrian Studies on Otto Neurath and the Vienna Circle. Partly translated from German. 1991 ISBN 0-7923-1276-7
134. W.R. Woodward and R.S. Cohen (eds.): *World Views and Scientific Discipline Formation.* Science Studies in the [former] German Democratic Republic. Partly translated from German by W.R. Woodward. 1991 ISBN 0-7923-1286-4
135. P. Zambelli: *The Speculum Astronomiae and Its Enigma.* Astrology, Theology and Science in Albertus Magnus and His Contemporaries. 1992 ISBN 0-7923-1380-1
136. P. Petitjean, C. Jami and A.M. Moulin (eds.): *Science and Empires.* Historical Studies about Scientific Development and European Expansion. ISBN 0-7923-1518-9
137. W.A. Wallace: *Galileo's Logic of Discovery and Proof.* The Background, Content, and Use of His Appropriated Treatises on Aristotle's *Posterior Analytics.* 1992 ISBN 0-7923-1577-4
138. W.A.Wallace: *Galileo's Logical Treatises.* A Translation, with Notes and Commentary, of His Appropriated Latin Questions on Aristotle's *Posterior Analytics.* 1992 ISBN 0-7923-1578-2
Set (137 + 138) ISBN 0-7923-1579-0

Boston Studies in the Philosophy of Science

139. M.J. Nye, J.L. Richards and R.H. Stuewer (eds.): *The Invention of Physical Science.* Intersections of Mathematics, Theology and Natural Philosophy since the Seventeenth Century. Essays in Honor of Erwin N. Hiebert. 1992 ISBN 0-7923-1753-X
140. G. Corsi, M.L. dalla Chiara and G.C. Ghirardi (eds.): *Bridging the Gap: Philosophy, Mathematics and Physics.* Lectures on the Foundations of Science. 1992
ISBN 0-7923-1761-0
141. C.-H. Lin and D. Fu (eds.): *Philosophy and Conceptual History of Science in Taiwan*. 1992
ISBN 0-7923-1766-1
142. S. Sarkar (ed.): *The Founders of Evolutionary Genetics.* A Centenary Reappraisal. 1992
ISBN 0-7923-1777-7
143. J. Blackmore (ed.): *Ernst Mach – A Deeper Look.* Documents and New Perspectives. 1992
ISBN 0-7923-1853-6
144. P. Kroes and M. Bakker (eds.): *Technological Development and Science in the Industrial Age.* New Perspectives on the Science–Technology Relationship. 1992 ISBN 0-7923-1898-6
145. S. Amsterdamski: *Between History and Method.* Disputes about the Rationality of Science. 1992 ISBN 0-7923-1941-9
146. E. Ullmann-Margalit (ed.): *The Scientific Enterprise.* The Bar-Hillel Colloquium: Studies in History, Philosophy, and Sociology of Science, Volume 4. 1992 ISBN 0-7923-1992-3
147. L. Embree (ed.): *Metaarchaeology*. Reflections by Archaeologists and Philosophers. 1992
ISBN 0-7923-2023-9
148. S. French and H. Kamminga (eds.): *Correspondence, Invariance and Heuristics.* Essays in Honour of Heinz Post. 1993 ISBN 0-7923-2085-9
149. M. Bunzl: *The Context of Explanation*. 1993 ISBN 0-7923-2153-7
150. I.B. Cohen (ed.): *The Natural Sciences and the Social Sciences.* Some Critical and Historical Perspectives. 1994 ISBN 0-7923-2223-1
151. K. Gavroglu, Y. Christianidis and E. Nicolaidis (eds.): *Trends in the Historiography of Science.* 1994 ISBN 0-7923-2255-X
152. S. Poggi and M. Bossi (eds.): *Romanticism in Science.* Science in Europe, 1790–1840. 1994
ISBN 0-7923-2336-X
153. J. Faye and H.J. Folse (eds.): *Niels Bohr and Contemporary Philosophy.* 1994
ISBN 0-7923-2378-5
154. C.C. Gould and R.S. Cohen (eds.): *Artifacts, Representations, and Social Practice.* Essays for Marx W. Wartofsky. 1994 ISBN 0-7923-2481-1
155. R.E. Butts: *Historical Pragmatics.* Philosophical Essays. 1993 ISBN 0-7923-2498-6
156. R. Rashed: *The Development of Arabic Mathematics: Between Arithmetic and Algebra.* Translated from French by A.F.W. Armstrong. 1994 ISBN 0-7923-2565-6
157. I. Szumilewicz-Lachman (ed.): *Zygmunt Zawirski: His Life and Work.* With Selected Writings on Time, Logic and the Methodology of Science. Translations by Feliks Lachman. Ed. by R.S. Cohen, with the assistance of B. Bergo. 1994 ISBN 0-7923-2566-4
158. S.N. Haq: *Names, Natures and Things.* The Alchemist J̄abir ibn Ḣayyān and His *Kitāb al-Aḣjār* (Book of Stones). 1994 ISBN 0-7923-2587-7
159. P. Plaass: *Kant's Theory of Natural Science*. Translation, Analytic Introduction and Commentary by Alfred E. and Maria G. Miller. 1994 ISBN 0-7923-2750-0
160. J. Misiek (ed.): *The Problem of Rationality in Science and its Philosophy.* On Popper vs. Polanyi. The Polish Conferences 1988–89. 1995 ISBN 0-7923-2925-2
161. I.C. Jarvie and N. Laor (eds.): *Critical Rationalism, Metaphysics and Science*. Essays for Joseph Agassi, Volume I. 1995 ISBN 0-7923-2960-0
162. I.C. Jarvie and N. Laor (eds.): *Critical Rationalism, the Social Sciences and the Humanities*. Essays for Joseph Agassi, Volume II. 1995 ISBN 0-7923-2961-9
Set (161–162) ISBN 0-7923-2962-7
163. K. Gavroglu, J. Stachel and M.W. Wartofsky (eds.): *Physics, Philosophy, and the Scientific Community.* Essays in the Philosophy and History of the Natural Sciences and Mathematics. In Honor of Robert S. Cohen. 1995 ISBN 0-7923-2988-0

164. K. Gavroglu, J. Stachel and M.W. Wartofsky (eds.): *Science, Politics and Social Practice.* Essays on Marxism and Science, Philosophy of Culture and the Social Sciences. In Honor of Robert S. Cohen. 1995 ISBN 0-7923-2989-9
165. K. Gavroglu, J. Stachel and M.W. Wartofsky (eds.): *Science, Mind and Art*. Essays on Science and the Humanistic Understanding in Art, Epistemology, Religion and Ethics. Essays in Honor of Robert S. Cohen. 1995 ISBN 0-7923-2990-2
Set (163–165) ISBN 0-7923-2991-0
166. K.H. Wolff: *Transformation in the Writing.* A Case of Surrender-and-Catch. 1995 ISBN 0-7923-3178-8
167. A.J. Kox and D.M. Siegel (eds.): *No Truth Except in the Details.* Essays in Honor of Martin J. Klein. 1995 ISBN 0-7923-3195-8
168. J. Blackmore: *Ludwig Boltzmann, His Later Life and Philosophy, 1900–1906.* Book One: A Documentary History. 1995 ISBN 0-7923-3231-8
169. R.S. Cohen, R. Hilpinen and R. Qiu (eds.): *Realism and Anti-Realism in the Philosophy of Science.* Beijing International Conference, 1992. 1996 ISBN 0-7923-3233-4
170. I. Kuçuradi and R.S. Cohen (eds.): *The Concept of Knowledge.* The Ankara Seminar. 1995 ISBN 0-7923-3241-5
171. M.A. Grodin (ed.): *Meta Medical Ethics*: The Philosophical Foundations of Bioethics. 1995 ISBN 0-7923-3344-6
172. S. Ramirez and R.S. Cohen (eds.): *Mexican Studies in the History and Philosophy of Science.* 1995 ISBN 0-7923-3462-0
173. C. Dilworth: *The Metaphysics of Science*. An Account of Modern Science in Terms of Principles, Laws and Theories. 1995 ISBN 0-7923-3693-3
174. J. Blackmore: *Ludwig Boltzmann, His Later Life and Philosophy, 1900–1906* Book Two: The Philosopher. 1995 ISBN 0-7923-3464-7
175. P. Damerow: *Abstraction and Representation.* Essays on the Cultural Evolution of Thinking. 1996 ISBN 0-7923-3816-2
176. M.S. Macrakis: *Scarcity's Ways: The Origins of Capital.* ACritical Essay on Thermodynamics, Statistical Mechanics and Economics. 1997 ISBN 0-7923-4760-9
177. M. Marion and R.S. Cohen (eds.): *Québec Studies in the Philosophy of Science.* Part I: Logic, Mathematics, Physics and History of Science. Essays in Honor of Hugues Leblanc. 1995 ISBN 0-7923-3559-7
178. M. Marion and R.S. Cohen (eds.): *Québec Studies in the Philosophy of Science.* Part II: Biology, Psychology, Cognitive Science and Economics. Essays in Honor of Hugues Leblanc. 1996 ISBN 0-7923-3560-0
Set (177–178) ISBN 0-7923-3561-9
179. Fan Dainian and R.S. Cohen (eds.): *Chinese Studies in the History and Philosophy of Science and Technology.* 1996 ISBN 0-7923-3463-9
180. P. Forman and J.M. Sánchez-Ron (eds.): *National Military Establishments and the Advancement of Science and Technology.* Studies in 20th Century History. 1996 ISBN 0-7923-3541-4
181. E.J. Post: *Quantum Reprogramming*. Ensembles and Single Systems: A Two-Tier Approach to Quantum Mechanics. 1995 ISBN 0-7923-3565-1
182. A.I. Tauber (ed.): *The Elusive Synthesis: Aesthetics and Science.* 1996 ISBN 0-7923-3904-5
183. S. Sarkar (ed.): *The Philosophy and History of Molecular Biology: New Perspectives.* 1996 ISBN 0-7923-3947-9
184. J.T. Cushing, A. Fine and S. Goldstein (eds.): *Bohmian Mechanics and Quantum Theory: An Appraisal.* 1996 ISBN 0-7923-4028-0
185. K. Michalski: *Logic and Time.* An Essay on Husserl's Theory of Meaning. 1996 ISBN 0-7923-4082-5
186. G. Munévar (ed.): *Spanish Studies in the Philosophy of Science.* 1996 ISBN 0-7923-4147-3
187. G. Schubring (ed.): *Hermann Günther Graßmann (1809–1877): Visionary Mathematician, Scientist and Neohumanist Scholar.* Papers from a Sesquicentennial Conference. 1996 ISBN 0-7923-4261-5

188. M. Bitbol: *Schrödinger's Philosophy of Quantum Mechanics.* 1996 ISBN 0-7923-4266-6
189. J. Faye, U. Scheffler and M. Urchs (eds.): *Perspectives on Time.* 1997 ISBN 0-7923-4330-1
190. K. Lehrer and J.C. Marek (eds.): *Austrian Philosophy Past and Present*. Essays in Honor of Rudolf Haller. 1996 ISBN 0-7923-4347-6
191. J.L. Lagrange: *Analytical Mechanics*. Translated and edited by Auguste Boissonade and Victor N. Vagliente. Translated from the *Mécanique Analytique, novelle édition* of 1811. 1997 ISBN 0-7923-4349-2
192. D. Ginev and R.S. Cohen (eds.): *Issues and Images in the Philosophy of Science.* Scientific and Philosophical Essays in Honour of Azarya Polikarov. 1997 ISBN 0-7923-4444-8
193. R.S. Cohen, M. Horne and J. Stachel (eds.): *Experimental Metaphysics.* Quantum Mechanical Studies for Abner Shimony, Volume One. 1997 ISBN 0-7923-4452-9
194. R.S. Cohen, M. Horne and J. Stachel (eds.): *Potentiality, Entanglement and Passion-at-a-Distance*. Quantum Mechanical Studies for Abner Shimony, Volume Two. 1997 ISBN 0-7923-4453-7; Set 0-7923-4454-5
195. R.S. Cohen and A.I. Tauber (eds.): *Philosophies of Nature: The Human Dimension.* 1997 ISBN 0-7923-4579-7
196. M. Otte and M. Panza (eds.): *Analysis and Synthesis in Mathematics.* History and Philosophy. 1997 ISBN 0-7923-4570-3
197. A. Denkel: *The Natural Background of Meaning*. 1999 ISBN 0-7923-5331-5
198. D. Baird, R.I.G. Hughes and A. Nordmann (eds.): *Heinrich Hertz: Classical Physicist, Modern Philosopher.* 1999 ISBN 0-7923-4653-X
199. A. Franklin: *Can That be Right?* Essays on Experiment, Evidence, and Science. 1999 ISBN 0-7923-5464-8
200. D. Raven, W. Krohn and R.S. Cohen (eds.): *The Social Origins of Modern Science.* 2000 ISBN 0-7923-6457-0
201. Reserved
202. Reserved
203. B. Babich and R.S. Cohen (eds.): *Nietzsche, Theories of Knowledge, and Critical Theory.* Nietzsche and the Sciences I. 1999 ISBN 0-7923-5742-6
204. B. Babich and R.S. Cohen (eds.): *Nietzsche, Epistemology, and Philosophy of Science.* Nietzsche and the Science II. 1999 ISBN 0-7923-5743-4
205. R. Hooykaas: *Fact, Faith and Fiction in the Development of Science*. The Gifford Lectures given in the University of St Andrews 1976. 1999 ISBN 0-7923-5774-4
206. M. Fehér, O. Kiss and L. Ropolyi (eds.): *Hermeneutics and Science.* 1999 ISBN 0-7923-5798-1
207. R.M. MacLeod (ed.): *Science and the Pacific War.* Science and Survival in the Pacific, 1939-1945. 1999 ISBN 0-7923-5851-1
208. I. Hanzel: *The Concept of Scientific Law in the Philosophy of Science and Epistemology.* A Study of Theoretical Reason. 1999 ISBN 0-7923-5852-X
209. G. Helm; R.J. Deltete (ed./transl.): *The Historical Development of Energetics.* 1999 ISBN 0-7923-5874-0
210. A. Orenstein and P. Kotatko (eds.): *Knowledge, Language and Logic.* Questions for Quine. 1999 ISBN 0-7923-5986-0
211. R.S. Cohen and H. Levine (eds.): *Maimonides and the Sciences.* 2000 ISBN 0-7923-6053-2
212. H. Gourko, D.I. Williamson and A.I. Tauber (eds.): *The Evolutionary Biology Papers of Elie Metchnikoff.* 2000 ISBN 0-7923-6067-2
213. S. D'Agostino: *A History of the Ideas of Theoretical Physics.* Essays on the Nineteenth and Twentieth Century Physics. 2000 ISBN 0-7923-6094-X
214. S. Lelas: *Science and Modernity.* Toward An Integral Theory of Science. 2000 ISBN 0-7923-6303-5
215. E. Agazzi and M. Pauri (eds.): *The Reality of the Unobservable.* Observability, Unobservability and Their Impact on the Issue of Scientific Realism. 2000 ISBN 0-7923-6311-6

216. P. Hoyningen-Huene and H. Sankey (eds.): *Incommensurability and Related Matters*. 2001 ISBN 0-7923-6989-0
217. A. Nieto-Galan: *Colouring Textiles*. A History of Natural Dyestuffs in Industrial Europe. 2001 ISBN 0-7923-7022-8
218. J. Blackmore, R. Itagaki and S. Tanaka (eds.): *Ernst Mach's Vienna 1895–1930*. Or Phenomenalism as Philosophy of Science. 2001 ISBN 0-7923-7122-4
219. R. Vihalemm (ed.): *Estonian Studies in the History and Philosophy of Science*. 2001 ISBN 0-7923-7189-5
220. W. Lefèvre (ed.): *Between Leibniz, Newton, and Kant.* Philosophy and Science in the Eighteenth Century. 2001 ISBN 0-7923-7198-4
221. T.F. Glick, M.Á. Puig-Samper and R. Ruiz (eds.): *The Reception of Darwinism in the Iberian World.* Spain, Spanish America and Brazil. 2001 ISBN 1-4020-0082-0
222. U. Klein (ed.): *Tools and Modes of Representation in the Laboratory Sciences*. 2001 ISBN 1-4020-0100-2
223. P. Duhem: *Mixture and Chemical Combination.* And Related Essays. Edited and translated, with an introduction, by Paul Needham. 2002 ISBN 1-4020-0232-7
224. J.C. Boudri: *What was Mechanical about Mechanics*. The Concept of Force Betweem Metaphysics and Mechanics from Newton to Lagrange. 2002 ISBN 1-4020-0233-5
225. B.E. Babich (ed.): *Hermeneutic Philosophy of Science, Van Gogh's Eyes, and God*. Essays in Honor of Patrick A. Heelan, S.J. 2002 ISBN 1-4020-0234-3
226. D. Davies Villemaire: *E.A. Burtt, Historian and Philosopher*. A Study of the Author of The Metaphysical Foundations of Modern Physical Science. 2002 ISBN 1-4020-0428-1
227. L.J. Cohen: *Knowledge and Language*. Selected Essays of L. Jonathan Cohen. Edited and with an introduction by James Logue. 2002 ISBN 1-4020-0474-5
228. G.E. Allen and R.M. MacLeod (eds.): *Science, History and Social Activism: A Tribute to Everett Mendelsohn*. 2002 ISBN 1-4020-0495-0
229. O. Gal: *Meanest Foundations and Nobler Superstructures*. Hooke, Newton and the "Compounding of the Celestiall Motions of the Planetts". 2002 ISBN 1-4020-0732-9
230. R. Nola: *Rescuing Reason*. A Critique of Anti-Rationalist Views of Science and Knowledge. 2003 Hb: iSBN 1-4020-1042-7; Pb ISBN 1-4020-1043-5
231. J. Agassi: *Science and Culture*. 2003 ISBN 1-4020-1156-3
232. M.C. Galavotti (ed.): *Observation and Experiment in the Natural and Social Science.* 2003 ISBN 1-4020-1251-9
233. A. Simões, A. Carneiro and M.P. Diogo (eds.): *Travels of Learning*. A Geography of Science in Europe. 2003 ISBN 1-4020-1259-4
234. A. Ashtekar, R. Cohen, D. Howard, J. Renn, S. Sarkar and A. Shimony (eds.): *Revisiting the Foundations of Relativistic Physics*. Festschrift in Honor of John Stachel. 2003 ISBN 1-4020-1284-5
235. R.P. Farell: *Feyerabend and Scientific Values*. Tightrope-Walking Rationality. 2003 ISBN 1-4020-1350-7
236. D. Ginev (ed.): *Bulgarian Studies in the Philosophy of Science*. 2003 ISBN 1-4020-1496-1
237. C. Sasaki: *Descartes Mathematical Thought*. 2003 ISBN 1-4020-1746-4
238. K. Chemla (ed.): *History of Science, History of Text*. 2004 ISBN 1-4020-2320-0
239. C.R. Palmerino and J.M.M.H. Thijssen (eds.): *The Reception of the Galilean Science of Motion in Seventeenth-Century Europe*. 2004 ISBN 1-4020-2454-1
240. J. Christianidis (ed.): *Classics in the History of Greek Mathematics*. 2004 ISBN 1-4020-0081-2
241. R.M. Brain and O. Knudsen (eds.): *Hans Christian Ørsted and the Romantic Quest for Unity.* Ideas, Disciplines, Practices. 2005 ISBN 1-4020-2979-9
242. D. Baird, E. Scerri and L. McIntyre (eds.): *Philosophy of Chemistry.* Synthesis of a New Discipline. 2005 ISBN 1-4020-3256-0

Boston Studies in the Philosophy of Science

243. D.B. Zilberman, H. Gourko and R.S. Cohen (eds.): *Analogy in Indian and Western Philosophical Thought.* 2006 ISBN 1-4020-3339-7
244. G. Irzik and G. Güzeldere (eds.): *Turkish Studies in the History and Philosophy of Science*. 2005 ISBN 1-4020-3332-X
245. H.E. Gruber and K. Bödeker (eds.): *Creativity, Psychology and the History of Science*. 2005 ISBN 1-4020-3491-1
246. M. Janssen, J. Renn, J. Norton, S. Tilmann and J. Stachel: *General Relativity in the Making*. Einstein's Zurich Notebook, Introduction and Source. 2006 ISBN 1-4020-3940-9
247. D. Ginev: *The Context of Constitution*. Beyond the Edge of Epistemological Justification. 2006 ISBN 1-4020-4712-6
248. J. Renn and K. Gavroglu (eds.): *Positioning the History of Science.* 2007 ISBN 1-4020-5419-X
249. M. Schemmel: *The English Galileo*. Thomas Harriot's Work on Motion as an Example of Preclassical Mechanics. 2007 ISBN 1-4020-5498-X
250. M. Janssen, J.D. Norton, J. Renn, T. Sauer and J. Stachel: *The Genesis of General Relativity*. Sources and Interpretations. 2007 ISBN 1-4020-3999-9
251. V.F. Hendricks, K.F. Jørgensen, J. Lützen and S.A. Pedersen (eds.): *Interactions*. Mathematics, Physics and Philosophy. 2006 ISBN 1-4020-5194-8
252. J. Persson and P. Ylikoski (eds.): *Rethinking Explanation*. 2007 ISBN 1-4020-5580-3
253. J. Agassi: *Science and History*. A Reassessment of the Historiography of Science. 2008 ISBN 1-4020-5631-4
254. W.R. Laird and S. Roux (eds.): Mechanics and Natural Philosophy before the Scientific Revolution. 2007 ISBN 1-4020-5966-7
255. L. Soler, H. Sankey, P. Hoyningen-Huere (eds.): *Rethinking Scientific Change and Theory Comparison*: Stabilities, Ruptures, Incommensurabilities? 2007. ISBN 1-4020-6274-2
256. W. Spohn: *Causation, Coherence and Concepts*. A Collection of Essays. 2008 ISBN 1-4020-5473-0
257. Yet to be published
258. M.J. Futch: *Leibniz's Metaphysics of Time and Space*. 2008 ISBN 1-4020-8236-8

Also of interest:
R.S. Cohen and M.W. Wartofsky (eds.): *A Portrait of Twenty-Five Years Boston Colloquia for the Philosophy of Science, 1960-1985.* 1985 ISBN Pb 90-277-1971-3
Previous volumes are still available.