INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
Clusters' last stand: Toward a theory of the process of meaning-making in science

Chokr, Nader N., Ph.D.

Rice University, 1991
NOTE TO USERS

Page(s) not included in the original manuscript are unavailable from the author or university. The manuscript was microfilmed as received.

394

This reproduction is the best copy available
RICE UNIVERSITY

CLUSTERS' LAST STAND:

Toward a Theory of the Process of Meaning-Making in Science

by

Nader N. CHOKR

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
DOCTOR IN PHILOSOPHY

APPROVED, THESIS COMMITTEE

[Signatures]

Richard E. Grandy, Director
Professor of Philosophy

C. Kenneth Waters, Asst.
Professor of Philosophy

Sydney M. Lamb, Professor of
Linguistics and Semiotics

Houston, Texas

April, 1991
CLUSTERS’ LAST STAND:
Toward a Theory of the Process of Meaning-Making in Science
Nader N. CHOKR

Abstract
The nature of the process of meaning-making in science has been one of the central problems in the philosophy of science of the 20th century. Yet, in spite of strenuous efforts by many able philosophers and historians of science over the past three decades or so, our understanding of this process continues to be unsatisfactory and fragmented at best. The need for an adequate account has been particularly exacerbated by the "infamous" and often misinterpreted problem of incommensurability (of meaning), and its alleged consequence, the incomparability of scientific theories --which presumably threatens the rationality, objectivity, and progress of science.

In this project, I argue that a new and revised cluster theory can be articulated, which meets the objections typically raised against (i) traditional (contextual or cluster) theories of meaning [Carnap, Kuhn, Gasking, Putnam, Achinstein] and (ii) theories of reference [Scheffler, Putnam, Kitcher]. Such a theory is not only based on more plausible assumptions and principles, but, in addition, it satisfies the main adequacy requirements formulated by proponents of a "cognitive-historical approach" [Shapere, Nersessian, Kuhn]. I am thus concerned not just with refuting "the entering wedge" of the argument against a defense of cluster theory, but with offering a relatively developed theory, sufficiently fleshed out to permit appreciation of its distinctiveness and evaluation of its merits.

I argue that the new cluster theory provides not only an adequate account of the process of meaning-making in science, but also a nuanced and context-sensitive one, which exhibits the fine-structure of the history of science. It is thus capable of accounting for the different kinds and degrees of meaning and reference changes in science. Furthermore, when applied in a case-study of the "chemical revolution," it accounts for that which has escaped most philosophical theories, namely, radical conceptual change without discontinuity, or even, as result of (and simply within) a broader framework of continuous conceptual change. The new cluster theory constitutes a proposal showing how the comparability of scientific theories is possible, how we have in fact been comparing them all along, despite "local incommensurabilities" of various kinds and degrees. Such a theory offers new insights into the developments of the chemical revolution in particular, but also into the structure and process of scientific revolutions in general. In short, it gives us a new framework for understanding the rationality, objectivity, and progress of science.
ACKNOWLEDGEMENTS

I wish to thank Richard Grandy, Kenneth Waters, and Sydney Lamb (Department of Linguistics), the members of my dissertation committee, for their generous and balanced guidance. Without it, this project would not have appeared in its present form. I also would like to express my deepest gratitude to all the faculty members of the Department of Philosophy, who have taught me how to be a better student of philosophy. Especially, I would like to acknowledge my debt to Richard Grandy, my advisor, for the probing questions and challenges that he confronted me with throughout my tenure at Rice University, and in particular, during the laborious years of research and writing which have resulted in this manuscript.
CONTENTS

INTRODUCTION

Part I: MEANING APPROACHES

1 Philosophical and Historical Context 24
   [A] From Frege to Carnap 24
   [B] Contextual Theories of Meaning 26
   [D] The Real Problem/Theory Comparison 38

2 Kuhn [Postscript, SSR 1970] 41

3 Concept-Representations: Sets vs. Clusters 58

4 Traditional Versions of Cluster Theories 69
   [A] Gasking 69
   [B] Putnam 80
   [C] Achinstein 92

Part II: REFERENCE APPROACHES 112

1 Scheffler's Focus on Reference 112

2 Putnam's Causal Theory of Reference 120

3 Kitcher's Reference Potential Approach 145

Part III: RECENT DEVELOPMENTS IN PHILOSOPHY OF SCIENCE 178

1 Shapere's Transtheoretical Approach 182

2 Nersessian's Cognitive-Historical Approach 207

3 Kuhn Revisited 233
Part IV: A NEW AND REVISED CLUSTER THEORY 268

1 A New Cluster Theory 269
   [A] Motivation and Context 269
   [B] Representational Problems 302
   [C] Developmental Problems 322

2 Case-Study: "The Chemical Revolution" 338
   [A] Representational Problems 341
   [B] Developmental Problems 368

3 Residual Problems 381

NOTES 384

BIBLIOGRAPHY 426
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSR</td>
<td>Structure of Scientific Revolutions</td>
</tr>
<tr>
<td>CTM</td>
<td>Contextual Theories of Meaning</td>
</tr>
<tr>
<td>CTR</td>
<td>Causal Theory of Reference</td>
</tr>
<tr>
<td>CST</td>
<td>Context-Sensitive Theories</td>
</tr>
<tr>
<td>CIT</td>
<td>Context-Insensitive Theories</td>
</tr>
<tr>
<td>CORC's</td>
<td>Chain of Reasoning Connections</td>
</tr>
<tr>
<td>Cp</td>
<td>Cluster of properties (term-type)</td>
</tr>
<tr>
<td>cp</td>
<td>cluster of properties (term-token)</td>
</tr>
<tr>
<td>r's</td>
<td>referent(s)</td>
</tr>
<tr>
<td>FRP</td>
<td>Field of Referential Possibilities</td>
</tr>
<tr>
<td>OT</td>
<td>Oxygen Theory</td>
</tr>
<tr>
<td>PT</td>
<td>Phlogiston Theory</td>
</tr>
</tbody>
</table>
LIST OF TABLES AND FIGURES

Double-Language Model [p.15]
Traditional Contextual Theories of Meaning [p.29]
Kitcher's Table: Phlogiston Theory/Modern Theory [p.149]
Shapere's Model [p.161]
Nersessian's Concept-Type: "Electromagnetic Field" [p.192]
Kuhn's Recent Views [p.228]
Philosophical/Historical Context [pp.236-7]
Traditional Cluster Theories [p.247]
New Cluster Theory [p.251]
Principles of the New Cluster Theory [p.258]
Shapere/Nersessian Model [pp.265]
Shapere/Nersessian Model (Improved Version) [p.265]
New Cluster Model [pp.266]
Fragment of a Cluster Network <Minerals> [p.275]
Summary: Case-Study of "Chemical Revolution" [pp.316-7]
Mind/Language/World [p.332].
Quand la philosophie est pratiquée
dans la conscience même de ses limites.
INTRODUCTION

The nature of the process of meaning-making in science has been one of the central problems in the philosophy of science of the 20th century. Yet, in spite of strenuous efforts by many able philosophers (and historians) of science over the past three decades or so, our understanding of this process continues to be unsatisfactory and fragmented at best.1 The need for an adequate account or theory of this process has been particularly exacerbated by the 'infamous' and often misinterpreted "problem of incommensurability (of meaning)," and its alleged consequence, the incomparability of scientific theories—which threatens presumably the rationality, objectivity, and progress of science.

In accord with some of the most recent developments in philosophy of science, which urge us quite rightly to regard science itself as a (cognitive and historical) process [Giere, 1980; Hull, 1980; Nersessian, 1987; Rubinstein et al., 1984],2 I propose to focus on the process of meaning-making in science, and to articulate a more satisfactory and adequate (cluster) theory of this process. Throughout this project, I shall thus be concerned only with those aspects of the process of meaning-making in science which are related or relevant to questions of comparison and evaluation of scientific theories. Some of these questions are: How do scientific terms (or concepts) acquire
meanings'? Where should we locate the locus of meaning (in the individual? or in the community? or elsewhere?) How should 'meanings' be represented at some point in time, over time? What is the nature of the relationship between 'meanings' and 'reference'? How do these 'meanings' change --as a result of "theory-change" or change in "background context"? What cognitive mechanisms and processes are involved? How do these meaning-changes constitute "radical meaning-variance" and lead subsequently to extreme and radical "incommensurability"? How does incommensurability of 'meanings' affect reference -determinations? How does it affect the comparison of scientific theories? Does it imply necessarily incomparability of scientific theories --as it has often been presumed? Does it imply therefore strong and sceptical relativism, and even irrationalism in science? That there can be no rational and objective account of the progress of science? Of the "evolution and continuity in scientific change" (Shapere, 1989), i.e., of evolution and continuity despite discontinuities of various kinds and degrees in scientific development?

The contemporary debate about theory-comparison and theory-choice has often been presented as a conflict between two clusters of ideas. On the one hand, we have the notions of rationality and progress in science, objectivity (both ontological and epistemological). On the other hand, we
have the view that 'meanings' are theory-dependent, with associated claims about the theory-dependence of observations, the "contextual theory of meaning," incommensurability and incomparability to which such a theory is presumed to lead inevitably. I intend to show that contrary to what many philosophers have assumed, there is no real conflict or contest between these two clusters of ideas,5 that they are in fact compatible.

To cut a long story short, I will show that this kind of conclusion, of conflict between these two clusters of ideas, is reached only because of inadequate theoretical and methodological assumptions. More positively however, I will argue that the cluster theory that I envisage is in fact a new form of "contextual theory of meaning" with a new cluster of theoretical and methodological assumptions --perhaps so new and different that it can no longer be considered as such, i.e., as a cluster theory; perhaps so new and different that it can be considered as such only by name (?:).6 In any case, according to this theory, the 'meaning' of a scientific term (or concept) is best characterized in terms of a cluster of (correlated and assertible) properties, to be understood only with respect to a given multivariate and variable "background context", relative to a given scientific community. Accordingly, it will vary in various ways with changes in "background
context." And, as I will show throughout and in Part IV in particular, this is quite consistent with a broadly rational and objectivist (in the epistemological sense) account of science and scientific development. Note that to the extent that I can, I will leave aside the issue of realism and anti-realism, or objectivity (in the ontological sense). Rather than this being a deficiency or failure, I think it could in fact translate into a comparative theoretical advantage for the cluster theory that I have in mind, if it can be shown in the end to accommodate either position.7 Furthermore, even though it would have been appropriate in some sense to devote a chapter to Quine's thesis of indeterminacy --a possible source of incommensurability, I will postpone such a project for another time and place, and thus keep the present one limited in scope and manageable. In the meantime, it shall suffice to take stock of my view on the subject.8 If Quine's thesis is correct, then we would indeed be faced with a "dramatic source of incommensurability" (Newton-Smith, 1981, p.180). In fact, it would not even make sense to suppose that we can come to glimpse, let alone understand (even partially), what previous scientists meant or referred to by the terms within their theories. This would be a far more radical and extreme consequence than Kuhn (or for that matter Feyerabend) ever contemplated with the notion of 'incommensurability'. 
How seriously should we then take Quine's thesis and its consequence? How conclusive is it? In view of the numerous, sometimes even contradictory, characterizations over a period of 30 years of the thesis of underdetermination (of theory by evidence), on which the thesis of indeterminacy is based, I would argue that it is hard to take such a thesis seriously, at least in its most extreme form. In fact, Quine's latest comments on this subject seem to justify such a position (1975; 1981).

The various accounts proposed by a number of philosophers and historians of science continue to be at odds with each other, as each of these accounts has its own list of adequacy requirements, and as each is plagued by difficulties and problems of its own. They exemplify as I see it three major kinds of approaches, which I shall call for my present purposes:

[I] Meaning Approaches (also called "Descriptivist Approaches," and which include contextual theories of meaning and traditional versions of cluster theory;

[II] Reference Approaches; and, for he most recent developments in philosophy of science on the subject here of interest, [III] Cognitive-Historical Approaches. 9

The motivating factor of this project is that a cluster theory of the process of meaning-making in science has not thus far been considered seriously enough. It has been too
quickly dismissed as inadequate on the basis of some of the common difficulties and objections which have been raised against traditional versions of this kind of theory (that is, based on the notion of cluster). In this project, I attempt to make the best possible case for a new and revised cluster theory.10 Such a theory (in the loose sense) is motivated essentially by the realization that many, perhaps most, scientific concepts are not definable in terms of necessary and/or sufficient conditions or even in terms of a cluster of properties or descriptions, "enough" of which must be satisfied for the corresponding terms to be applicable. By incorporating a number of new principles and by drawing on recent contributions in cognitive science, history of science, sociology of science and linguistics, this new cluster theory will constitute, I claim, an important step toward a fuller and more detailed account of the process of meaning-making in science. I do not wish to claim that such a theory is complete and entirely satisfactory or even definitive. Rather, as suggested above, I wish to show that the case for a cluster theory has never really been made, and that it can in fact be improved and revised so as to take care of the common difficulties and objections raised against it.11 Furthermore, it can be shown to be based on reasonable and plausible (theoretical and methodological) assumptions, and to satisfy most, if not
all, of the requirements made over the years for an adequate theory of the process of meaning-making in science. Finally, I will show that it has, as a theory exemplifying in a new and different way the Meaning Approach, a number of selective advantages, theoretical and methodological, over a theory exemplifying the Reference Approach, its principal competitor, or even over a theory exemplifying the Cognitive-Historical Approach—at least, as conceived thus far.12

The detailed plan of this project is as follows. In Part I, I present and critically examine what I called the "Meaning Approaches." After a brief background on Frege's legacy, I consider (in chapter 1) the "contextual theory of meaning" as it is found in Hempel [1950-2], Carnap [1956] and Kuhn [1962],13 and discuss the central problem which this theory has presumably led to concerning the interpretation of science, namely extreme and radical incommensurability (of meaning). In this respect, my point is that Kuhn's notion and thesis of incommensurability has often been misinterpreted, partly due to Kuhn's terminological confusions and partly because many commentators have simply misunderstood Kuhn's point. In any case, it is important to keep in mind, as a reconsideration of Kuhn's Postscript in SSR, 1970 [modulo Grandy, 1983; (chapter 2)] shows, that there are in fact different sources
and kinds of incommensurability; and that none of them (except perhaps one) raises the kind of threat envisaged by pessimistic "incommensurabilists" about the rationality and objectivity of science, and in particular, about the comparability of scientific theories. One should also note that most of the accounts which have been proposed over the years have been determined or else constrained at least in part by their respective (assumed or explicit) construed of the notion and thesis of incommensurability. And so in a sense, the evaluation of the various accounts of interest in this whole project will depend at least in part on their respective construed and formulation. Having said this, I do not deny that there is a real and serious problem for philosophers of science to deal with: how to account for the fact that we can and do compare different scientific theories, even across revolutionary divides --without raising the kinds of problems that the contextual theory of meaning engenders? As I show in chapter 1 [B] [C] and 3, traditional contextual theories of meaning (including cluster theories) fail to provide such an account mainly because of their assumed infectious holism, essentialism of meaning, and inadequate concept- or meaning-representation. In chapter 4, I finally consider and evaluate three specific traditional versions of cluster theory (Gasking [1960], Putnam [1962] and Achinstein [1968]). I argue in short
that these traditional versions also fail in part because they are still essentialist in their assumptions, unclear about questions of epistemic access and linguistic competence, because they are based on inadequate or incomplete principles, respectively of omnifocality, over-determination, and semantical relevance.

In Part II, I turn to the "Reference Approaches." I begin in chapter 1 by looking at Scheffler's argument in [1967], that in order to avoid and evade problems of "incommensurability," we need to shift our attention from meanings (or senses), which change more or less radically as a result of theory change, and to focus on reference, which, he presumed, remained stable and unchanging, thus guaranteeing continuity and commensurability in scientific change. In chapter 2, I examine Kripke/Putnam's causal theory of reference [1972; 1980/1973; 1975], which has been considered by many as solving or dissolving the incommensurability problem (e.g., Hacking, 1983, pp.74–91). Far from solving or dissolving the problem here in question, I argue that it faces in fact many difficulties and problems, not the least of which are created by the "rigidity and essentialism of reference" that it posits, and its inability to deal satisfactorily with certain crucial episodes in the development of science which involve not only meaning-changes but also reference-changes, sometimes
so radical that certain scientific terms which used to refer to a given entity or type of entity are later found to refer to an entirely new and different entity or type of entity, or even not to refer to anything at all. In light of these and other theoretical and methodological difficulties faced by Putnam’s causal theory of reference, Kitcher has attempted to formulate an improved version of such a theory, in terms of reference potential (chapter 3). I examine this new and improved version of CTR and evaluate it. My assessment is that it faces some other or additional difficulties of their own, even though it represents clearly a significant and laudable effort to "rehabilitate" Putnam’s CTR. Bearing in mind some of the contributions and insights of these various versions of a CTR, I then turn to some more recent developments in philosophy of science.

In Part III, I consider the "Cognitive-Historical Approaches" of Shapere [1984, 1989], Nersessian [1984, 1987; 1989], and Kuhn [1982-1990]. Even though their accounts fall under this general rubric, Shapere develops more specifically what I shall call a "Reason-based Transtheoretical Approach," while Nersessian articulates properly speaking a "Cognitive-Historical Approach." As for the most recent Kuhn, he develops what I shall call a [locally holistic] Verification Theory of Meaning" based on the notions of "lexicon's structure" or "lexical field" as
the locus of meaning. I review and evaluate these different, yet very similar, approaches, with an eye toward the new and revised cluster theory that I wish to propose. For, in my assessment, Kuhn, Shapere and Nersessian are in fact "closet cluster theorists," who deny that they are because, I think, they only have in mind a traditional version of cluster theory.15

It is fair to say that they all assume that incommensurability does not necessarily nor always lead to incomparability. When and if incommensurability obtains, it is often "local incommensurability." Their discussion makes it clear that incommensurability is not as disastrous and threatening as most philosophers have construed it to be. Taking this as a starting-point, they each propose a number of minimal requirements for an adequate theory of the process of meaning-making in science. Essentially, such an theory should show that while we may encounter incommensurability, "local incommensurability" that is, there is somehow a greater degree of commensurability, and how to account for it. Rather denying that we can compare scientific theories, or the difficulties of doing so at times, these authors argue quite rightly that what we need to ask is how it is possible to compare theories, and go on to make a proposal to answer this question.

Finally in Part IV, on the basis of the contributions and
insights of all previous accounts, I articulate the new and revised cluster theory, its substantive claims and responses to traditional objections, its assumptions (both theoretical and methodological), the adequacy requirements that it satisfies, the basic principles (structural and interpretative) on which it based (chapter 1). Then (in chapter 2), by way of beginning its evaluation, I turn to a case-study of the "Chemical Revolution" --which led to the demise of "phlogiston theory" and its replacement with Lavoisier's oxygen theory. Having discussed the merits and advantages of the new cluster theory, I close this part by considering (in chapter 3) some residual problems of this theory, outlining thus a few areas for possible future research.

In presenting and critically examining the various accounts mentioned, I follow roughly the same strategy as the one sketched above. I articulate their respective substantive claims (theoretical and methodological), point out their assumptions, discuss their merits and underscore their limits and difficulties. Also, insofar as it is appropriate I attempt to bring out their respective underlying conception of philosophy of science.1

In this last respect, it might be helpful to situate briefly the approach that I take in this project within the recent history of the philosophy of science, and to
characterize its underlying conception. Since roughly the late 50’s and early 60’s, the philosophy of science has been undergoing a major transformation. This process began when the positivists’ view of science and scientific knowledge was challenged as one with little resemblance to and little relevance for the understanding of real science, of actual scientific practices by real scientists. Later on, further challenges and objections were raised against other aspects of the positivists’ view (observation, theory, meaning, reference, contexts, explanation, etc), to the point where it is hard to find today anyone who would unreservedly support the positivists’ view. And yet, this view continues to have an influence on today’s philosophy of science, in particular over its problems and methods of analysis. Perhaps the best way to understand the development which took place is to examine two imaginary extreme positions, situated at both ends of a continuum, which functions as the Scylla and Charybdis between which most philosophers of science seek to articulate their views. My characterization of these extreme positions is not meant to reflect the specific views of any individual philosopher, even though, as we shall see, some philosophers and historians of science come close to espousing one or the other position. I present these two positions as ideal-types, each very tempting and each untenable in
extreme form. Most philosophers of science, including myself, fall somewhere in between.

At one extreme, we have "autonomism," the position, which positivists (e.g., early Carnap, Hempel) have either espoused at some point or come close to adopt, that philosophy of science is (ought to be) an "autonomous" discipline, carried from a prescriptive "external" standpoint (epistemological, metaphysical, or logical) in order to draw the warrant for its account of the nature of science and scientific knowledge. Accordingly, it is assumed that there is some sort of first philosophy or philosophical methodology from which, by some sort of analytic, a priori, or transcendental techniques philosophers can derive, dictate, or legitimate the standards by which scientific work should be evaluated. It is also assumed that even though the application of such standards to particular explanations, hypotheses, or theories requires substantive knowledge of the actual procedures and practices followed in science, one is supposed to be able to derive, establish and justify the standards abstractly or rather independently of particular scientific knowledge. To establish and justify standards is a philosophical, not a scientific enterprise. (Leplin, 1988; Burian, 1987).

At the other extreme, we have "historicism," the position, which Kuhn, Feyerabend [and Shapere (?)] have either
espoused or come close to espouse, according to which philosophy of science should rely for its warrant upon a careful "internal" description of actual scientific procedures and practices, of how scientists actually proceed, or have in the past proceeded. The business of philosophy of science does not consist in prescribing the standards (abstractly) by which scientific work should be evaluated. The function of different elements (e.g., concepts, explanations, hypotheses, laws, theories) are not studied in the abstract, but in the practices of scientists themselves. The tasks of philosophers of science might be (i) to explicitly and perspicuously articulate the concepts, theories, methods, and standards employed by actual scientists (a task which even the best scientists often find difficult, according to McMullin [1970]); (ii) to sort out the different kinds of contexts in which different standards are employed consistently, and thus viewed as "appropriate"; and (iii) to evaluate and dismiss those confused interpretations which result from applying standards "appropriate" in one context to another, for example, by those who "whiggishly" judge past theories by present standards. (Kuhn, 1986; Burian, 1987). It is important to note with respect to (ii) that pure historicists do not claim to have the sort of independent authority which could establish that the standards in question should be employed
in that context. In other words, historicists claim to describe standards and the contexts in which they are employed consistently, but they do not prescribe or independently evaluate the standards they describe (see McMullin, 1979).

Given these kinds of considerations, the autonomist approach to philosophy of science views science as the ideal form of human knowledge, and supposes that its nature can best be understood by beginning from a general philosophical theory of knowledge. It also takes it to have a logical structure of demonstration and justification, and supposes that it can be judged just as one would any other formal or rigorously defined linguistic system. The goals of science are defined in terms of maximizing justified true belief about the world or in terms of maximizing the three S's, scope, simplicity, and systematicity of our experientially grounded knowledge, expressed in theories (i.e., sets of statements). Accordingly, the history of science is itself considered, approximately, as the continuous accumulation of "hard facts" and of ever exact, comprehensive and systematically "unified" theories covering those facts. Since it is also widely believed that the acceptance of a theory is not justified unless, on the available evidence, it is highly probable that the theory is at least approximately true, it is often held that successful
theories are never false and only rarely abandoned completely. On such a view, superseded theories are presumably subsumed under their successors. On general grounds, it is held that (in principle) the scientific knowledge at any given time can be stated in a "unified" language consisting of a theory-neutral ("observational") base and various theoretical sub-languages built upon that base. Nersessian summarizes the "double-language model" implied by this view (as in Table 1 adapted from 1984, pp.9-12, esp.10):

Table 1: Double-Language Model

<table>
<thead>
<tr>
<th>Standard Account</th>
<th>Experience</th>
<th>C</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>T &lt;--------</td>
<td>O</td>
<td></td>
</tr>
</tbody>
</table>

Legend: C= Conceptual; G= Given of experience; T= Theoretical; O= Observational; CS= Conceptual structure; Th= Theory; Arrow: direction of flow of empirical meaning.

In the autonomist tradition, philosophy of science has consisted simply in applied epistemology. The standards which should be employed in judging the adequacy and truth of factual claims, the degree of support for hypotheses and theories, and the justification for and value of theories,
are prior to the truths of science. They are derived for the most part, but not all, from philosophical theories of knowledge [cf: Hooker, 1975]. The primary tool used to deal with such matters is logical (or linguistic) analysis of the (idealized) language of science. Such a tool enables one presumably to articulate the structure of sound) scientific concepts, laws, hypotheses, confirmations, explanations and theories and to compare "rationally reconstructed" concepts, laws, hypotheses, explanations, theories, etc, against the resultant norms of scientific practices. Such "rational reconstructions" of the "finished products" of science at a particular stage of its development constitute presumably the proper and most perspicuous objects of philosophy of science. The history (or psychology or sociology) of science is of little or no concern for practitioners of such philosophy of science, even though they cannot be wholly unconcerned about serious divergences between their account of the nature of science and the course science has actually followed in particular cases. Typically, they draw upon "case-studies" in the history of science for counter-examples, illustrations and/or indirect support of philosophical theories.

Turning now to those philosophers of science who have taken the "historicist turn." They may view science as an activity seeking to maximize justified belief about the
world. However, they treat the effective goals and standards of judgement and evaluation in science, even the nature of science, as varying with time, undergoing shifts with larger and major theories and traditions. Generally, they view the course of science as involving two quite different phases. One the one hand, there is the phase of problem-solving activities, of self-corrective accumulations of facts and development of theories which have already been formulated more or less fully (called "normal science" by Kuhn). On the other hand, there is the phase which is characterized by (Kuhn's "scientific revolutions") radical changes of "world-view," "paradigm" (Kuhn), "highest-level background theory" (Feyerabend), "research programmes" (Lakatos). In this case, there may be loss of facts and explanatory power, change of standards governing acceptance of fact-claims and theories, loss of the ability to compare rival theories in a straightforward and uncomplicated way. These difficulties come about as a result of conceptual changes, which undermine the hope of working in a theory-neutral (observational) language, unifiable to cover the whole body of science. In order for observations to succeed in testing a theory, they must be reported in a language compatible with the concepts or categories of that theory. But since differing fundamental theories employ presumably incompatible or incommensurable concepts or
categories, the languages in which observations germane to these theories are reported cannot be adequately mapped onto each other. Or so the extreme historicists would have us believe.

In the historicist tradition, philosophy of science can never be the "mere" application of epistemology, or what has been suggested in more recent times, of a theory of language. For, what counts as adequate means of introducing a new concept or term, of supporting hypotheses is something learned, in part, from the study of science and its history. "We have learned how to learn" by studying science and its history (Shapere, 1984). Whatever standards independent of science which are advanced by epistemologists, are too weak to accomplish the evaluation of scientific claims or theories unless they are further specified or revised on the basis of the relevant science. The source of the authoritative standards is, in good part at least, internal to the "paradigms," or "high-level background theories," "research programmes" found in science. Such standards cannot be judged by "pure" philosophy. Normatively, the most that a philosopher can do presumably is to formulate and apply the rational standards for acceptance, development, correction, and improvement of ongoing science implicit within science itself. The primary tool for such a task is a historical and cognitive analysis of the actual
formation of scientific concepts, conceptual frameworks, theories, etc. Such an analysis is unreliable if it is restricted to the "finished products" of science (not to mention the rational reconstructions thereof). These tend to obscure major aspects of the development and evolution of the scientific enterprise, its underlying mechanisms and processes (cognitive and otherwise). Kuhn pointed in this respect (1970, chap.xi) that, as "repositories of finished products," science textbooks falsify the history of science. According to McMullin (1970), the history of science is relevant for two reasons: (i) because it provides complete and extensive case-studies, of a kind one could not recover from looking at contemporary science; (ii) more fundamentally, because it allows one to study science in its all-important temporal dimension, and to take into account the developmental aspect of science, the characteristic ways in which a concept or a theory is modified, altered, and eventually replaced.

More recently however, in the past ten or fifteen years, there has emerged a new approach within the discipline called the "New Philosophy of Science." Such an approach seeks to transform or reformulate the problems and methods of philosophy of science by charting a different course, perhaps somewhere between the two extreme positions of autonomism and historicism, or by attempting to draw the
proper measure and significance of the "historicist turn". Even though there is still no consensus on this subject, most philosophers of science today, according to Nersessian would make the following contentions:

(1) Science is an open-ended, on-going activity, whose character has changed significantly in its history.
(2) Science is not a monolithic enterprise.
(3) Good science can lead to false theories.
(4) Science has its roots in everyday circumstances, needs, methods, concepts, etc, of human beings.
(5) Through examination of actual science we will learn how to understand such notions as 'observation', 'theory', 'meaning', 'reference', 'explanation', 'progress', and 'rationality'.
(6) The boundary between a 'scientific' question and a philosophical question, especially as regards foundational problems, is often blurry.
(7) The function of the philosophy of science is descriptive, critical, and normative.
(8) The philosophy of science is not self-sufficient, insofar as the methods of philosophy need to be supplemented by concepts, principles, theories, etc, drawn from other disciplines; though it remains to be determined just what can be used and how (1987, p.vii).

The proliferation in recent times of "case-studies" of past and present scientific practices seems to lend support to these tenets. Case-studies seem to be more than simply means to develop counter-examples, criticisms, and/or illustrations of philosophical theories. They seem to suggest that "the study of the processes by which theories and concepts, among other things, are formulated and accepted or rejected is a prerequisite to philosophical understanding " (Nersessian, 1987, p.viii; italics and underlining added). Accordingly, the task of the "New Philosophy of Science" should be to attempt to deal with the
existing complexities of actual scientific practices, rather than create neat and sweeping abstract theories. But then, as Nersessian quite rightly puts it, the question is: how can (or should) philosophers of science do this? By what methods can they legitimately study the processes of scientific inquiry and its products? And one might add: By what methods can they legitimately study the process of meaning-making in science?

Having said that, it would be interesting as we go through this project, and examine the views of a number of philosophers of science, to see where they each fit on the continuum between autonomism and historicism, or else, whether and how they have articulated the "historicist turn." It would also be interesting to see where the new and revised cluster theory fits, and to what extent (if it does), it exemplifies the new approach to philosophy of science.
PART I: MEANING APPROACHES

Chapter 1: Philosophical and Historical Context

A. Frege's Legacy

As the distinction between sense ('meaning' or intension) and reference (extension) looms large in this project, it is only appropriate to begin by explaining it. This distinction was first made properly clear by Frege (1892, 1970). He showed how in the theory of language we need to separate the independent entity that a word stands for (its 'reference') from the aspect of that word's use which determines it to have that reference (its 'sense'). In a manner of speaking, we can conceive of the sense of word as that which directs us to the thing or things for which it stands. However, as several contemporary philosophers pointed out (e.g., Kripke, 1972, p.277n28; 1980; Kitcher, 1978, p.543), Frege specified the notion of sense in two different ways: (i) the sense of a term or expression is "that which is grasped by someone who understands the term or expression," and it is also (ii) the way in which the reference is determined, or else, "the manner in which the reference is presented." Also, as Frege made clear, with the well-known example of "Morning Star" and "Evening Star," it is perfectly possible for two words with different senses to have the same reference: "names with different senses but the same reference correspond to different routes leading to the same destination" (Dummett,
1973, p.96).

In the context of science which is of interest to us here, this latter point would translate as follows: a scientific term (e.g., 'mass') can have different senses (or meanings) --in two different theories (Newtonian and Einsteinian), and yet still have the same reference and (be used to) stand for the same quantity. Supposing that the senses or meanings of scientific terms might depend somehow on their theoretical contexts, and so might vary and change with theoretical changes, is it plausible to argue that the references of scientific terms remain always the same and do not vary at all with theoretical changes? Granted that it is plausible enough that the former does indeed occur, does it on the other hand follow that the references will change as well with theoretical changes? Not necessarily. In fact, one could argue that "it is surely implausible to suppose that any general argument (...) for variance in reference could be given. Lack of identity of reference of the same terms used in different theoretical contexts must be established in each case by particular argument for each context" (Martin, 1971, pp.19-20).2

Frege's views have unquestionably influenced and even shaped the theories of meaning formulated by the early logical positivists (in the 1920-40's).3 In fact, they have been so influential that a number of recent philosophers
have argue that problems and paradoxes (of meaning-variance) arise in our current theories of meaning only because insufficient attention has been paid to Frege's views, and in particular to the distinction between sense and reference. Whatever is the case, we shall be in a better position to determine at a later point. In the meantime, it must noted that for the early positivists, like Frege, the fundamental unit of meaning is the term or statement. According to their theory of meaning, known as a "verification theory of meaning," the meaning of a term or expression is "its method of verification," i.e., the method determining the necessary and/or sufficient conditions of application of the term or expression.4

B. Traditional Contextual Theories of Meaning

However, by the 1950's, under repeated criticisms of the verification principle itself, the logical positivists moved from a "naive" verificationism to a more sophisticated, yet still problematic, one. They abandoned the idea that the fundamental unit of meaning is the term or statement, and adopted instead the view that theoretical (or linguistic) context determines the meanings of individual terms or expressions. This is, I take it, what Carnap, one of the most influential positivists, stated in his famous 1956 work:

The definition of meaningfulness must be relative to a
theory T, because the same term may be meaningful with respect to one theory but meaningless with respect to another. It is clear that the definition (of a theoretical term) must be relative to T, because the question whether a certain term in L of theory T is significant cannot possibly be decided without taking into consideration the postulates by which it is introduced (1956, pp. 49-50).

For my present purposes, it does not seem necessary to review the long history of self-criticisms and incremental refinements which has led such prominent figures of the positivist tradition to this sort of view. It is this sort of view which has been referred to as the "contextual (or network) theory of meaning." During the past three decades or so, it has gained wider acceptance in one form or another beyond the confines of positivism. I will not herein concern myself with its origins and developments, I will only confine my attention to some of the problems presumably generated by such a view in particular (i.e., Kuhn's) concerning the interpretation of science and to how these problems can be overcome or dissolved, whatever the case may be.

Incidentally, it is significant to note, what most commentators overlooked or disregarded: the views on meaning developed by Kuhn (and Feyerabend, for that matter) in the 50's -60's were not, at least in this respect, different from those that the positivists (e.g., late Carnap, 1956), their archfoes, were already defending. [The first edition of Kuhn's Structure of Scientific Revolutions, 1962 (hereafter SSR), was initially published as part of the
positivist series on the *International Encyclopedia of Unified Science* co-edited by Neurath, Carnap and Morris, vol. 2/Number 21.6. And so the positivists have their share of responsibility for whatever difficulties and problems the contextual theory of meaning has led to (for further substantiation of this point, see Jane English, 1978; Newton-Smith, 1981; Graham Oddie, 1989).

Three sets of issues regarding contextual theories of meaning and their ancestors must be raised at this point.

(I) What is to be considered the meaning-determining context? And how determining or determinant is it? Is it to be the "theory" in which the term occurs? In that case, would not the sense of "context" be as vague and ambiguous as that of "theory"? Would it not assume that theories form some sort of easily recognizable entity, easily individuated and distinguished from one another? And that it is perfectly clear to determine what constitutes part of a theory and what does not? Is it instead a "special type of theory"? But then, what kind of theory is it exactly? Is it a "high-level background theory," as Feyerabend suggested? Or is it something broader than a theory? How much broader than a theory? What would be the cut-off point? Is it instead, as Kuhn [1962] proposed, a "paradigm"? Is it as Kuhn [1970] proposed, a "disciplinary matrix"? Or is it even broader and much more all-encompassing? It
seems that all these different versions of the contextual theory of meaning have their difficulties and problems. To focus however on Kuhn's version which is based on the notion of "paradigm," it has been shown repeatedly and quite convincingly, as we shall see, that such a notion is vague and ambiguous, and yet at the same time still all too-determinant (Mastermann, 1970; Shapere, 1971, 1984). As such, it has been argued that it is the principal source of the difficulties that Kuhn's version of contextual theory (at least the early one) encounters with regards to the interpretation of science. However these various versions of the contextual theory are construed ultimately and developed subsequently in greater detail, it is fair in the meantime to say that they all concentrate on the "meaning-determining context" as something which consists of a large number of inter-related beliefs, assertions, and factual claims. We might at later point find a more suitable and appropriate way to characterize this context and the extent to which it is determinant, or simply shapes the meaning of the terms or expressions embedded in it.

(II) Insofar as concepts and concept-formation and development constitutes a central aspect of scientific theories and research, we must inquire about the role and place of concepts and concept formation in the contextual accounts of meaning and meaning-change here in question. As
Nersessian correctly points out:

The creation of concepts through which to comprehend, structure, and communicate about physical phenomena constitutes much of the scientific enterprise. Concepts play a central role in the construction and testing of the laws and principles of a theory. The introduction of new concepts and/or the alteration of existing ones is a crucial step in most changes of theory. And in many scientific controversies what is at issue is disagreement over the interpretation of fundamental concepts. In short, articulating concepts is a central aspect of scientific research. Thus, our understanding of science is seriously deficient if we fail to examine the question of how scientific concepts emerge and are subsequently altered (1987, p.161; italics added).

Yet, such examinations have been given very little attention, if any, in the positivist theories. This may explain in part why these accounts have failed. To use the well-known distinction codified by Reichenbach, the study of actual concept formation and development was deemed to belong to the "context of discovery" which is of interest to the historian and the psychologist, and not to the "context of justification" which is considered to be the proper province of the philosopher. Thus, when Hempel wrote his Fundamentals of Concept Formation in Empirical Science (1952), the object of his study was not the actual formation of scientific concepts, but rather the "rationally reconstructed" conceptual structure of science. Within such a framework, concept formation in science found to be unproblematic and was characterized as continuous and cumulative, with the new or altered
conceptions being simply logical extensions of previous concepts. Accordingly, change, continuity, and comparison of meaning was considered unproblematic. It was assumed that the observational language of science provides a "theory-neutral" basis to which all theoretical terms can ultimately be reduced. However, as it is now widely accepted, the major problem with such a framework was that of specifying the nature of the observation/theoretical distinction and of the reduction of theoretical terms to observational ones.

By the late 50's and early 60's however, it became increasingly clear that the meaning of terms was determined by their "theoretical context," that the distinction between "contexts" (discovery vs. justification) was questionable, and furthermore, that no sharp and clear distinction can be drawn between the observational and theoretical languages of a scientific theory. With these realizations, "the problem of meaning-change came to the fore, as the \textit{bete noire} of post-positivistic philosophy" (Nersessian, 1987, p.162). It seemed only logical to conclude therefore that there can be no "neutral" point of comparison for the terms of a theory and its competitor or successor. According to Nersessian (1987, p.162), this was the view buttressed by the claims of Feyerabend and Kuhn that analysis of the history of science reveals 'incommensurability' of meaning to be a fact. And
subsequently, change of meaning has come to be seen as the result of a catastrophic 'revolution', as taking place in a discontinuous manner and in such a way that the concepts of the new theory completely replace those of the previous one. In this respect, one important question must however be raised: does the fact that there is no "neutral" (observational) language for comparing two different scientific theories mean that no comparison whatsoever is possible — e.g., in terms of other suitable considerations, or by some other route, via for example a meta-language.9 (cf: Grandy’s point, 1983; Part I, chap.2)?

(III) Most contextual theories (including Kuhn’s early account) seem to have shared with the early positivist theories their conception of the "meaning" of a term: it consists of the conditions of application of the term. As mentioned earlier, these traditional theories have drawn, after Frege, a distinction between the sense (or "meaning") of a term and its reference. They have also held that the former determines the latter. The sense or "meaning" of a term is then taken to consist of conditions of application of the term, while the reference is considered to be the entity or entities of which it is true. The conditions of applications of the term were formulated in terms of descriptions: they first considered to be a set of necessary and sufficient conditions, then,
a set of necessary or sufficient conditions, and finally, a "cluster" of conditions "enough" of which must be satisfied for the term in question to be applicable. With the contextual theories, even though it is now the (theoretical) context which is said to determine what the conditions of applications are. The meaning itself (of a term) is still considered to consist of such conditions, however they are characterized, in terms of sets or in terms of clusters. What I mean to suggest here in anticipation is that the problems and difficulties of contextual theories may be due or depend at least to a great extent on their conception of the meaning of a term, and consequently, on their representation of the concept corresponding to it. I shall have more to say on this point in forthcoming sections. In the meantime, the traditional (contextual or cluster) theories of meaning can be represented schematically as follows:

Background
Context (Theoretical) \(\xrightarrow{\text{determines}}\)

\[\begin{array}{c}
\text{Term: Meaning} \quad \xrightarrow{\text{Referent(s)}} \\
\text{[logically necessary \&/or]} \\
\text{[sufficient conditions]} \\
\text{[associated with term]} \\
\text{(monolithic)} \\
\text{(determinant)} \quad \text{[cluster of properties]} \\
\text{[associated with term]} \\
\text{[on linguistic grounds]}
\end{array}\]

C. The Problem of Incommensurability
Briefly, let us see how Kuhn's contextual theory of meaning leads presumably to a critical problem in the interpretation of science. According to Kuhn (1962), normal science is governed by an overarching "paradigm." The meanings of scientific terms are determined by the theory in which they occur, and thus by the paradigm in which this theory is formulated. On this view, even the same terms in different theories, different paradigmatic traditions, can be "incommensurable" and thus "incomparable." This incommensurability and incomparability is not limited to the meaning of terms, but most commentators have seized upon this aspect as the most seriously threatening (Giere, 1988, p.34).10 As Kuhn put it (1962, p.102), a paradigm is "the source of the methods, problem-field, and standards of solution accepted by any mature scientific community at any given time." In other parts of SSR, Kuhn extends this theory/paradigm dependence to "facts", "observations", to any alleged "given" in experience. The idea here being, if all these elements are also paradigm-dependent and determined, then they are incommensurable" and presumably "incomparable." From all this, it follows that "the normal scientific tradition that emerges from a scientific revolution is not only incompatible, but often actually incommensurable with that which has gone before" (1962, p.102). Even though Kuhn has not always been clear or
forthcoming about the justification of the "often" that he interposed in the statement above, and even though he has left us with no explicit way or technique to compare different theories/paradigms, it is important to note its significance (pace and contra Shapere, 1984 and other commentators). In a sense, it announces what Kuhn will later keep repeating and hammering upon. Contrary to what was believed under "positivist spells," there may be incommensurabilities (of various kinds, with different causes, and of different degrees). This may threaten or restrict our ability to compare straightforwardly two different theories. But does this mean that there is no commensurability at all in scientific development, that comparability is forever and always foreclosed, and finally that communicability across a revolutionary change in science is impossible? It seems to me that Kuhn never really thought so, even though his terminological vagueness and ambiguity led us to believe that he did. He later clarified his view on the subject (Postscript, SSR, 1970; 1982; 1988; see also Part I, chap. 2; Part III, chap. 3 for further discussion).

However, it is in this context, that is, via the extreme and radical incommensurability thesis which most commentators attributed to Kuhn (1962), that the problem of comparing different scientific contexts (theories or
something broader like a paradigm or a disciplinary matrix) became a central problem in the philosophy of science. If meanings (as well as facts, methods, standards, and everything else) are dependent and determined by a paradigm, a disciplinary matrix, or by a theoretical context, and if furthermore they are characterized as mentioned above, then in two different such contexts, the terms will have different meanings. The question then is: how can we compare the meanings, theories, paradigms or disciplinary matrices? Can they be compared at all?

At least two different views are possible. (i) The view which has been attributed to Kuhn (and Feyerabend) is one which denies that comparability is at all possible, and which, for this purpose, emphasizes the differences in the "meanings" of terms (e.g., "mass") in two different theories/paradigms (e.g., Newtonian and Einteinian physics), while relegating any similarities to a lesser status, as if they could not possibly constitute some basis for comparison. This view was taken to lead to the thesis of extreme incommensurability. Is this thesis the consequence of an examination of actual science, of the actual process of meaning-making in science, as presumably Kuhn argued? Or is it the consequence of an ill-founded and inadequate conception of meaning? —A conception in which meaning is determined by paradigmatic or theoretical context, and which
excludes similarities from the meaning of terms and retains only differences, and is thus based on an undefended and indefensible essentialism of meaning?

(ii) The other view is one which takes seriously the idea of degree of similarity of meanings, and claims that the latter could serve as a basis for (at least a limited) comparison of the meanings of terms in two different theories. On this view, one could then show that the term "mass" in Newtonian and Einsteinian physics have some degree of similarity of meaning, and thus are not totally incommensurable or incomparable. This kind of view would enable us to better make sense of the actual process of meaning-making in science. It would also make possible and make sense of the tasks of both historians and philosophers of science concerned with understanding and interpreting this process.

Let’s consider for a moment the thesis of extreme and radical incommensurability on its own merits, regardless of Kuhn’s or anybody else’s arguments for it. What comes to mind immediately is a question inspired by a Davidsonian line of attack (1984, pp.183-198): How can any two things be completely and utterly incommensurable and therefore incomparable? Along the same line, how can two theories, contexts, paradigms, languages or conceptual schemes be completely and utterly incommensurable and incomparable? The thesis of extreme and radical incommensurability seems
fundamentally flawed and clearly incoherent. If two scientific theories/contexts were utterly and completely incomparable as the thesis suggests, it would not be possible to make sense of the fact they are both called "scientific," or "theories" or even "contexts" (paradigmatic or otherwise); it would not be possible to say they differ in their standards, for example, or any other respects for that matter (methods, problems/solutions, domains and conceptions of explanation, etc). And yet, historians and philosophers of science, Kuhn and Feyerabend included, have made and still continue to make such comparisons, which visibly are across different paradigmatic or theoretical contexts (cf: Kitcher, 1978, pp.519-20; Shapere, 1984).11

D. The Real Problem/Theory-Comparison

I have just argued that the thesis of extreme and radical incommensurability/incomparability is incoherent and must be rejected. This does not mean however that we do not encounter "cases of incommensurability" and that there is no problem whatsoever about comparing scientific theories. Theses "cases," as I mentioned earlier, may be of different kinds, due to different causes, and of various degrees. Most often however there is only a "local and limited kind of incommensurability" and comparability is not totally foreclosed, even though it may at times be seriously hampered. The comparison of scientific theories however is
a real problem and must be acknowledged as such (see Part III and IV for further discussion). As we have learned from the history of science, scientific theories talk of very different sorts of entities, and even what the theories are about seems to be different. If two scientific theories, or contexts, differ in these and many other respects, as they certainly seem to do, on what basis do we compare them? Instead of concluding that we cannot, on the basis of some ill-conception of meaning as characterized above, we need to ask how we can do so and how in fact we do make such comparisons? Can we always compare straightforwardly without any loss in our ability? Do we always have good reasons for making comparisons? These questions should be viewed I think as an integral part of any adequate theory of the process of meaning-making in science.12 They have to do ultimately with the rationality and progress of science. Another important question, related yet prior to the ones just mentioned, that any adequate theory must address has to do with whether there is some connection between two scientific theories, or contexts, such that one may be said to be an ancestor of the second, and such that there is a traceable line of descent from the first to the second. This question concerns the historical evolution (and continuity) of science, while the previous one(s) concern the evaluation (and comparison) of scientific theories. As
such, the former is prior to the latter and must be settled before the latter can be dealt with. I will come back to this question throughout.

In the meantime, it shall suffice to take stock of Kuhn’s comments concerning these questions (Postscript, SSR [1970]). He asks us to "imagine an evolutionary tree representing the development of the modern scientific specialities from their common origins in, say, primitive natural philosophy and the crafts." And he then states that "a line drawn up that tree, never doubling back, from the trunk to the tip of some branch would trace a succession of theories related by descent (1970, p.205; italics added).

He goes on to add:

Considering any two such theories, chosen from points not too near their origin, it should be easy to design a list of criteria that would enable an uncommitted observer to distinguish the earlier from the more recent theory time after time. Among the most useful would be: accuracy of prediction, particularly of quantitative prediction; the balance between esoteric and everyday subject matter; and the number of different problems solved. Less useful for this purpose, though also important determinants of scientific life, would be such values as simplicity, scope, and compatibility with other specialities. Those lists are not yet the ones required, but I have no doubt that they can be completed. If they can, then scientific development is, like biological, a unidirectional and irreversible process. Later scientific theories are better than earlier ones for solving puzzles in the often quite different environments to which they are applied (1970,p.206; italics added).
Chapter 2: Kuhn’s Postscript to SSR (1970)

As we have seen, Kuhn’s contextual account favored a more holistic approach to meaning and rejected the reductionist view based on the observation/theoretical distinction, among others. As a result, one of the theses he has presumably argued for is the radical "meaning-variance thesis" and its consequence the infamous "incommensurability problem." Its essential implication is that it is not possible to compare two comprehensive theories T1 and T2 within the same scientific field, particularly if they are separated by a revolution, with respect to the meaning of all of their respective (theoretical) terms.1 Thus, he has argued, for example, that two scientific theories such as T1 = "phlogiston theory" and T2 = Lavoisier's "oxygen theory" are incommensurable and therefore incomparable. Similarly, he has argued that Ptolemaic and Copernican astronomy and that Newtonian dynamics and Einstein's special theory of relativity are incommensurable and incomparable (Kuhn, 1962). Presumably, he arrived at such "paradoxical" conclusions as a result of arguments of the following form: The meaning of a scientific term is to be viewed rather holistically in terms of the theory in which it is embedded. Its meaning is thus determined by the context of beliefs, values, practices and problems prevailing at some point in time for some scientist
or group of scientists. The meaning of a term in one scientific theory T1 is not the same as its meaning in some prima facie conflicting theory T2. No statement, in particular no observational statement, can contradict a statement containing the term in another theory. No observation statement which belongs to one theory can be used as an independent test for another theory. There is no observational or theory-neutral basis for comparison. Furthermore, there is no shared set of methodological rules (inductive logic). There is no external, common language which scientific theories shared which could serve as a basis for comparison. Therefore, there is no principled way nor objective basis to decide between scientific theories.

In support of such arguments, Kuhn has often pointed out certain kinds of situations from the history of science, which do appear to threaten the possibility of an objective comparison between rival scientific theories. These situations involve typically a special type of conceptual change, i.e., a radical referential change, of the key terms of a scientific theory, which culminates in an inability of the proponents of two rival theories to specify the referents of the terms used in presenting their respective theories.2

These kinds of situations must be taken seriously, but if the theses, in support of which they are mentioned, are
correct, then some important part of the historian's and philosopher's enterprise is impossible. More generally, we will not be able to express the content of past theories in modern (or later) terms. This raises the problem of intertranslatability between the old and the new language in a given scientific field traversed by a "revolution." As we have seen, the three problems of incommensurability, intertranslatability and incomparability are often taken to be inter-related.3

My goal ultimately is to show that a sensitive reading of certain episodes of the history of science, coupled with a more sophisticated approach to semantical issues (i.e., a new and revised version of cluster theory as I envisage it in Part IV) will yield instead the following theses: (i) While there may be considerable meaning-variance in science, there are degrees of meaning-variance, and there must be a far greater degree of commensurability. (ii) Even if there is incommensurability at times, i.e., of a limited kind and locally, it is not as disastrous as it has often been depicted. I will argue that an adequate cluster theory offers us a strategy for better understanding the graded character of meaning-change in science, and as a result, the process of meaning-making. Rather than simply denying that opposing theories within a scientific field traversed by a revolution, are incomparable and incommensurable, it will
attend to the possibility of formulating them in their respective language and comparing them.

Many proposals have been made by philosophers of different persuasion to deal with the "incommensurabilist" theses often attributed to Kuhn. This is not the time or space to discuss and evaluate them all systematically. However, one in particular, which has been overlooked in the literature, is worth singling out, and that is Grandy's (1983). Briefly put, it offers a perspicuous and rather illuminating reading of Kuhn's SSR (1962). A discussion of this proposal will serve among other things to clarify the often misinterpreted account made by Kuhn, and to show in what sense he already held (in 1970) the views mentioned above.

Grandy argues essentially for two conclusions: (I) there are different kinds and causes or sources of incommensurability and shows via a logical approach that (II) the problem of incommensurability (of meaning) is not as disastrous as it has been depicted. I fully agree with Grandy's analysis leading up to conclusion (I) and (II). However, as a proponent of a new and revised cluster theory, I suggest that we reach conclusion (II) via a historical/cognitive/linguistic approach anchored in actual scientific practices. This may go to show that these two approaches are not after all incompatible and could even be
complementary. In fact, one could well argue that the former approach (i.e., Grandy's) clears up, so to speak, the condition of possibility of the latter. I will first discuss (I) and then (II) in this order.

Interestingly enough, Grandy begins his investigation of the problem of incommensurability by proposing quite correctly that Kuhn's SSR (1962) be re-written in light of the new notion of the Postscript (1970). While the central element in the former was the inflationary notion of "paradigm," the central notion in the latter is that of "disciplinary matrix." "Disciplinary" because it refers to the common possession of the practitioners of a particular discipline; "matrix" because it is composed of ordered elements of various sorts, each requiring further specification (Kuhn, 1970, p.182). Thus, the incommensurability problem can and should no longer be viewed as between paradigms, but rather as between disciplinary matrices, or more precisely, some or all of their respective components. There are indeed several components to a disciplinary matrix. Some of these are (see Kuhn, 1970, pp.182-90; Grandy, 1983, pp.8-11): (1) Symbolic generalizations --which typically serve to define technical terms as well as their relations, and have to do with meaning (e.g., (i) $F = ma$; (ii) $S \rightarrow R$; (iii) metal + air$\rightarrow$ metal oxide + air poor in oxygen). (2) Metaphysical
assumptions --about the a priori or necessary features of the world, which constitute a constraint on the type of theory a scientist may find acceptable (e.g., Materialism vs Mechanism; Cartesian mechanics vs Newtonian mechanics). (3) Models (mathematical and otherwise, e.g., "gas- molecules = tiny billiard balls in random motion"; "electric circuit = a steady-state hydrodynamic system") --applicable to certain objects or processes, most often analogical, suggesting measurements, connections with observable phenomena, etc, which serve as a guide to developing the explicit scientific theory. (4) Values --concerning predictions (accuracy, quantitative/qualitative character); concerning simplicity and generality of theories, their compatibility with other current theories, the degree of explicitness of formulation of a theory, the relative ease of reproducibility of the results, etc. (5) Instruments --which are more often than is recognized associated with important changes in the meaning of scientific terms, in the perception of the world, in the problems to be solved, in the solutions thought to be acceptable (e.g., sealed glass retort; calorimeter; Wheatstone bridge; telescope; microscope; air-pump, etc). And finally, we have, in Kuhn's terms, "the most novel and least understood" constituent of a disciplinary matrix, namely, (6) examplars, i.e., shared examples, by way of which scientists acquire similar "similarity relations," or
else, the ability to group objects, events and entities into clusters based on such similarity relations. They include:
(a) particular types of solutions to problems that a scientist is exposed to in training or research; (b) high points of scientific achievement which influence scientists;
(c) fully specific instances of actual lab investigation (Kuhn, 1970, p.187; Grandy, 1983, p.11). They serve to illustrate and direct in different ways (i.e., more or less directly) the work of research scientists (e.g., Newton’s derivation of the law of falling bodies and Kepler’s laws from the law of universal gravitation and other principles; Lavoisier’s experimental demonstrations of the composition of air and water). They also serve to characterize the scientific group and even normal science. According to Grandy:

The centrality and importance of exemplars derives from their role in binding together the other elements of the disciplinary matrix in a way that provides specific indications of how research can proceed. One of the deep and controversial issues surrounding Kuhn’s work is whether the implicit guidance of exemplars can or should be rationally reconstructed in terms of rules” (1983,p.11; italics added).

Before coming back to this aspect of exemplars, which Grandy regards rightly as one of the sources of incommensurability, and other key terms of Kuhn’s work which must be revised, let us reconsider for a moment the incommensurability problem. Once again, I agree with Grandy when he concludes:
Once we have distinguished the separate elements of a disciplinary matrix, we can see that there are distinct kinds of incommensurability potentially associated with each element. The greater the number of components that change in the disciplinary matrix, the greater the number of kinds of incommensurability that arise" (Grandy, 1983, pp.14-5).

Thus, we could have corresponding to the different components of a disciplinary matrix, different kinds of incommensurability, and thus different degrees of incommensurability:

(1') Incommensurability of meaning — since "a difference in symbolic generalizations will often mean a difference in definitions, and thus in meaning" (Grandy, 1983, p.8).

Presumably, there was such an incommensurability between Priestley's phlogiston theory [PT] and Lavoisier's oxygen theory [OT] in their respective descriptions of familiar reactions:

[PT]: Metal + air ---> Calx of metal + phlogisticated air.
heat

[OT]: Metal + air ---> metal oxide + air poor in oxygen.5
heat

(2') Incommensurability of metaphysics — e.g., the famous German resistance to Lavoisier's theory is one good example. Another is the Cartesian resistance to Newtonian mechanics.

(3') Incommensurability of models — e.g., between Faraday's and Maxwell's respective model of the structure of the "electromagnetic field."

(4') Incommensurability of values. Kuhn writes in this
respect: "Though values are widely shared by scientists and though commitment to them is both deep and constitutive of science," the application of values can vary considerably, particularly "when the members of a particular community must identify crisis, or later, choose between incompatible ways of practising their discipline" (Kuhn, 1970, pp.185;184). Furthermore, they are "sometimes considerably affected by features of individual personality and biography that differentiate the members of the group" (Ibid., p.185).

(5') Incommensurability associated with instruments --which is not often given the attention it deserves. And finally, as one might expect, we may encounter:

(6') Incommensurability of exemplars --as shared examples. "More than other sorts of components of the disciplinary matrix, differences between sets of exemplars provide the community fine-structure of science" (Kuhn, 1970, p.187).

Thus, in the end, it is more appropriate to say that there are various kinds and causes of incommensurability. And also, various degrees of incommensurability. A detailed analysis could show that none undermines the objective comparison of scientific theories, or as Grandy puts it "the objectivity of science and its progress" (1983, p.15). He writes further:

Almost all occurrences of incommensurability can be interpreted as incommensurability of disciplinary matrices....I have argued that the degrees of incommensurability vary greatly but that these differences
in the disciplinary matrices need not imply incomparability. Nor the impossibility of incompatibility" (Ibid., p. 22; see end of present section for significance of underlining).

In fact, according to the later Kuhn (1982, p. 2), it seems like no matter how much incommensurability there is in scientific development, it is most often only "local incommensurability," and thus, not as threatening as he presumably depicted it earlier. Kuhn clearly has come to embrace such a position, as the title of one his recent papers suggests ("Commensurability, Comparability, and Communicability," 1982). In anticipation, I should say that an adequate cluster theory would adopt a similar position.

In order to further clarify and update Kuhn's work in light of the 1970 Postscript, we now need to consider other key terms of SSR, 1962. (i) A scientific community is defined by the paradigm that its members share. (ii) A scientific revolution is defined as a change of paradigm. (iii) Normal science is defined as a period of scientific development guided by a single paradigm; [anomalies are defined as those unsolved puzzles which lead to a revolution]. (iv) The meaning of a scientific vocabulary in a given theory changes when a paradigm changes (see Grandy, 1983, p. 7).

As Grandy quite rightly points out (1983, pp. 12-3, 13-4, 18), these characterizations (i)-(iv) must now be understood as follows: (i') A scientific community is a group that shares (all or part of) a disciplinary matrix; thus, there can be different communities of different kinds, sizes and types, depending on how much their respective members share
as a group. To be more precise, Grandy defines a scientific community as "those scientists whose similarity judgements are more alike than they are like [those of] members of other groups" (1983, p.18). These similarity judgements are presumably acquired on the basis of certain clusters of examplars rather than others. This is why examplars are characterized as one the sources of incommensurability. According to Grandy, however, "The 'group similarity relation' is a useful fiction that provides a brief way of speaking about the diverse though rather similar relations belonging to the individual members of a group." And he adds:

Kuhn's point is that the group need not be defined by perfect agreement in their judgements, but only in terms of their relative ease of agreement in contrast with the difficulties of communicating with members of other groups" (Grandy, 1983, p.18; see on this point Kitcher, 1982 [Part II] and Kuhn, 1988 [Part III]).

(ii') A scientific revolution is a change of all or part of a disciplinary matrix. As Grandy remarks in this respect, "one immediate question is whether the change requires a total or only a partial change in the matrix."

The answer might be that a change is revolutionary for group G if it involves a change in the shared elements of the disciplinary matrix...Thus, whenever we speak of a change being revolutionary we must relativize the claim to a particular level of analysis of scientific groups. It is only when a change involves the most general and widely shared elements of the disciplinary matrix [that] one can speak of a revolution tout court. In fact, I would suggest that the terminology be changed and that "revolution" be reserved for large scale changes. Another term such as "reconceptualization" would be more appropriate for the types of change under discussion. Thus the main contrast would be between what was a reconceptualizing change as opposed to normal (continuous) change for a given scientific community. Revolutions would be large scale reconceptualizations (1983, p.13; italics in text).

Perhaps it is partially Kuhn's fault that most readers of his work focused their attention on revolutions, and as a result overlooked "the equally but less dramatic changes
that frequently occur on a smaller scale and which constitute most of scientific development" (Grandy, 1983, pp.13-4). This unfortunate consequence was arguably due to the strategy that Kuhn adopted, i.e., to emphasize the phenomena of gestalt switches (or "changes in vision/outlook/worldview") which occur in science, and thus illustrate those kinds of changes which could most effectively show the importance of large scale revolutionary changes. (See Kuhn's explanation for this in his 1980 lectures, Part III). [In anticipation, one should note that a cluster theory of the process of meaning-making in science as I envisage it (see Part IV) would take into account both kinds of changes]. As one might expect, Grandy concludes his discussion about revolutions and scientific changes as follows:

Thus, the fundamental point about revolutions is the kind of change involved, and I suspect that a better understanding of this kind of change is more likely to be forthcoming from the study of small scale reconceptualizations rather than large since the analyses of the former are less likely to be entangled in larger issues [i.e., with social and religious implications] (1983, p.14; italics in text).

The important point about our present discussion however is that just as there are various degrees of incommensurability, there are various degrees of revolutions --since incommensurabilities accompany revolutions.

Also, just as the recognition that scientific communities can be discerned at many different levels has
important implications for our understanding of a scientific revolution, the recognition that one of the components of a disciplinary matrix, i.e., exemplars, serves to define not only a direction of research but also the scientific group, has important implications for characterizing normal science. Thus, (iii') normal science is guided by a cluster of exemplars, some of which have an impact only on small subgroups of the community, while others have an impact on the larger groups. [Anomalies will be of different kinds, and will be perceived differently, as a source of crisis or not, by different groups and sub-groups].

Similarly, one should say that:

(iv) the meaning of the scientific vocabulary of a given theory changes when (all or parts of) the corresponding disciplinary matrix changes i.e., the disciplinary matrix which can be attributed to the theorists by virtue of their membership in a given scientific community. Thus, once again, there will be various kinds and degrees of meaning-change—depending on which and how many components of the disciplinary matrix have changed— which must be relativized to a particular level of analysis of scientific groups.

Turning now to the second conclusion (II) that I attributed to Grandy earlier, namely that the incommensurability problem is not as disastrous as it has
been depicted, let me reiterate that it is only in the approach taken to reach this conclusion that a cluster theory as I envisage it differs from Grandy's. Furthermore I agree with his characterization of the problem of incommensurability between scientific theories as "one of the most disputed and misunderstood claims of SSR" (Grandy, 1983, p.18). In accord with what Kuhn himself had already pointed out in his 1982 paper, Grandy reminds us of the originary meaning of "incommensurability" and of the straightforward analogy being drawn. He writes explicitly:

In mathematics, the term "incommensurable" has a clear and precise meaning that is being used very straightforwardly in the analogy between scientific theories. Two magnitudes are incommensurable if there is no pair of integers \( n \) and \( m \) such that the ratio of the first magnitude to the second is represented by \( n/m \). The most famous historical example of incommensurable magnitudes are the side and diagonal of a unit square. Since 2 is an irrational number, no pair of integers represent the ratio in question. This does not imply however that we cannot compare the magnitudes in question with as great a degree of precision as is desired; for any pair \( m \) and \( n \), we can determine whether \( m/n \) is greater or less than the ratio, all that the incommensurability proof shows is that \( m/n \) is never equal to the ratio. The claim that scientific theories from different disciplinary matrices are incommensurable is intended to make an equally strong claim, but no stronger. (1983, p.19; my underlining).

While traditionally philosophers of science have sought to find a theory-neutral basis for comparison of scientific theories in an observation language (a sublanguage, so to speak), Grandy proposes that we sidestep the so-called "incommensurability of meaning" for comparison in a meta-language, i.e., via theories of truth. His approach is
properly speaking logical (for further details, see Grandy, 1983, pp.19-21). He considers various senses of incompatibility, since there can be no comparison if we cannot determine incompatibility. There is no need to review them here. It shall suffice to indicate that it is de facto incompatibility, the broadest sense of incompatibility, which Grandy thinks is required for understanding how incommensurable scientific theories can be compared. As he puts it (1983, p.20):

T1 in L1 and T2 in L2 are de facto incompatible if there is a language M containing correct (possibly partial) truth theories D1 and D2 for L1 and L2 respectively such that for some set of true sentences K of M,

\[ K, D1, D2 \models \neg [\text{True} \ L1 ("T1") \ \& \ \text{True} \ L2 ("T2")]. \]

The strategy that Grandy proposes is a two-step strategy: first, we show that not both theories can be true by using all of the information at our disposal about both the theories and the world; second, we attempt to determine which, if either, is compatible with the facts. Such a strategy assumes that "a sufficiently neutral metalanguage can be found, partly because the two theories in question are attempting to describe a single world" (1983, p.23; my underlining). Grandy cautions however, "if distinct disciplinary matrices are describing different worlds then none of the earlier comparisons via truth theories make sense" (my underlining). He is referring here to the possibility that Kuhn's text may be hiding a much deeper
sense of incommensurability, one that cannot be explained as in previous cases as a mismatch between (components of) disciplinary matrices. This is why Grandy said earlier that "almost all occurrences of incommensurability can be interpreted as incommensurability of disciplinary matrices."

He was keeping the more recalcitrant case for the end. He points to a puzzling passage in which Kuhn admits: "In a sense I am unable to explicate further, the proponents of competing paradigms (or disciplinary matrices) practice their trade in different worlds" (1962, p.150; my underlining). As Grandy puts it, there seems indeed to be an ontological chasm yawning here (see also Kuhn, 1988 [Part III] for a clarification of what he meant). If this holds, then Grandy's proposal becomes pointless. In order to place Kuhn's statement in proper perspective, and to begin assessing its viability as a thesis, Grandy concludes quite pertinently:

The positivist conception of science presupposed that we were given boxes into which we sorted the objects of nature. Kuhn (and others who have attended to the history of science) have argued, as discussed above, that we must first manufacture the boxes, deciding on their size, shape and number. And sometimes in mid-sorting we must redesign the boxes. But the metaphysical doctrine darkly hinted at in the last section suggests that scientists in different disciplinary matrices are not only using different boxes, but are using different objects too (1983, p.23; my underlining).

Since I cannot herein undertake the detailed philosophical analysis which might be required to fully
assess the viability of this thesis, I would like to close this discussion by making a few anticipatory remarks about of an adequate cluster theory and its metaphysical assumptions. One of the distinctive features of the cluster theory that I envisage will be its "neutral" stance concerning the issue of realism/anti-realism. This will I think constitute an advantage in that it will be able to accommodate either position. In response to Grandy's way of formulating the issue, one might take a "constructive realist" perspective, and point out that we do not manufacture the boxes arbitrarily, our box-making is motivated and constrained by the real world (and, one should add, by our perceptual/cognitive capacities or limits). There is one world, but there may be several possible ways of sorting the objects of nature, some equivalent, some not, and possibly some neither. However, for the most part, there are significant similarities, to say the least, between our various ways of sorting things out (cf: Grandy's principle of humanity, 1973). These similarities are of course greater among members of the same scientific community than among members of different communities.
Chapter 3: Concept-Representations: Sets vs. Clusters

I suggested earlier that a prerequisite for an account of the process of meaning-making in science is an adequate characterization of what is to be included in "meaning," of what constitutes the "meaning" of scientific terms. I also suggested that the problems of "meaning-variance" and its consequence, the thesis of extreme and radical incommensurability to which contextual theories have presumably led to, are due in part at least to an inadequate conception of meaning, an inadequate concept-representation, based on an undefended and perhaps indefensible essentialism. In line with the latter, the meaning of a term is said to be determined by and to consist of the conditions of application of the term --necessary and/or sufficient conditions, or a cluster of conditions "enough" of which must be satisfied for the term to be applicable. I shall now turn to a more explicit discussion and evaluation of this characterization. This shall serve as a motivation and background for the forthcoming discussions. In anticipation, we should expect that an adequate account will provide not only an adequate meaning-representation (or concept-representation) at some point in time but that it will also take into consideration the open-ended nature of the process of meaning-making in science, and will therefore provide an account of the developmental aspects involved therein.
Once again, according to contextual theories, the meaning of a term is taken to be determined and constituted by the set or cluster of properties (conditions or descriptions) associated with it, while its reference is constituted by the things or entities of which this set or cluster is true. Thus the meaning of a term is given by specifying the set or cluster of properties that belong to the things or entities constituting the reference. This conception provided not only a criterion for something to belong to the reference—in the sense of necessary and/or sufficient conditions, or in the sense of a cluster of conditions, but also an identification procedure for whether something falls into the reference or not (cf: Putnam, "Meaning and Reference," 1973, pp.699-700). Hence, a concept is represented, at a given point in time, by a "definition": that is, by a set of conditions, each of which is necessary and/or all of which are jointly sufficient to define it, or else, by a cluster of conditions enough of which must be satisfied. Consequently, a given thing or entity is or is not in the reference of a concept depending on whether or not it has each and/or all of the properties or enough (many, perhaps most) of the cluster of properties specified as constituting the meaning of the concept in question.

Upon closer examination, this "definitional"
representation faces however some serious difficulties. First, it assumes that it is possible to state a set of necessary and/or sufficient conditions or even a cluster of conditions which "define" a particular concept. Furthermore, it assumes that the properties used in identifying a particular entity (i.e., a substance or a natural kind), and which are specified as constituting the meaning of the term in question belong essentially to the entity. Second, it does not take into account the nature of the process of meaning-making in science. As such, it makes it impossible to carry out an examination of how we get from one concept to another in competing scientific theories, and thus, contributes in some sense to the problem of incommensurability.

Concerning the first point, it is acknowledged nowadays --in part as a result of Wittgenstein's work (1953) and more recently, of Rosch's experimental studies (1975)-- that it is extremely difficult, if not outright impossible, to do what is assumed by the "definitional" approach with respect to ordinary concepts (e.g., game, fish, bird, etc). This is also true of (many, perhaps most) scientific concepts, except in the case of those (e.g., Newtonian system, Carnot cycle, reversible process, etc) which are introduced by explicit "definitions." But how often does this happen within the process of meaning-making in science? When and
if it does happen however, is not the "definition" sooner or later found not to be applicable to the entities that it used to refer to, and perhaps to be applicable to other entities. This is to be expected given the open-ended nature of science. Furthermore, to state a set of necessary and/or sufficient conditions, or a cluster of conditions, in order to "define" a concept, presupposes that one can distinguish between 'essential' and 'accidental' properties, i.e., between those which all instances of the concept must (always) have and those which most instances might (possibly) have. But can this be done? To take some examples, "flies" is a property of most birds, but it is not essential since some birds do not fly; similarly, "solid at room temperature" is a property of most metals, but it is not essential since mercury, for example, which is a metal, is liquid at room temperature. Similarly, "...has atomic number 29" is nowadays considered to be a property (an important and central one) of 'copper', but is it essential? In what sense? Could we say that it is essential — given the present state of scientific knowledge? If we could, then "...has atomic weight #?" was essential given the previous prevailing state of scientific knowledge. Is the term 'essential' used in a meaningful way in this case, i.e., when it is thus relativized? Perhaps, it makes sense to consider various "grades of essentialism" and to argue
that some grade(s) are compatible with the scientific process while others are not. But is not 'essentialism' of whatever grade simply unscientific? (pace and contra Leplin, 1988). Suppose however that it makes sense to distinguish three "grades of essentialism" as follows:

**Grade 1**: necessary and sufficient conditions or properties satisfied in all possible worlds.

**Grade 2**: necessary or sufficient conditions or properties satisfied in all possible worlds.

**Grade 3**: cluster of conditions or properties enough of which are satisfied in all possible worlds.

Does Grade 3 involve any less commitment to essentialism? Does not essentialism of whatever grade simply contradict the very open-ended nature of the scientific process? And should not an adequate conception of the process of meaning-making in science avoid first and foremost any form or grade of essentialism? The short answer in my view is that it should.

The traditional way of conceiving of meaning is, as mentioned earlier, saddled with an undefended and indefensible essentialism of meaning. The properties used originally in identifying an entity (a substance or a natural kind), or alternatively, which are specified as constituting the meaning of the corresponding term, need not (and in fact do not) belong essentially to the entity,
substance or natural kind. As the process of science takes place, these properties may (even all) be found later not to belong to the entity, substance or kind. An alternative, yet plausible scenario might also be that other entities, substances or natural kinds may later be found to have all the properties here question and yet not be that kind of entity, substance, or natural kind. In accord with the nature of the process of meaning-making in science, the same point can be made about any other later "definitions." Thus, the properties ascribed to entities, or kinds of entities, originally or subsequently, cannot be treated as the essential, unalterable meanings of the terms used to refer to those entities or kinds, as "defining properties" of these terms. Interestingly enough, Kripke and Putnam have argued along a similar line for why a theory based on essentialism of meaning must be rejected; only they went on, as we shall see (in Part II) to propose a theory based on an essentialism of reference. The question that comes to mind at this point is this: Is this the only alternative? Could not there be a contextual theory of meaning, which does not, so to speak, "sin by contrary excess," i.e., which does not assume such an essentialism of meaning? Could this be a cluster theory (like Gasking's, Putnam's, or Achinstein's, or even a new and revised version like the one I will envisage in Part IV)? My short answer is that it
could well be the latter.

Concerning the second point, let us assume that a given concept "X" in theory T is "defined" by a set of necessary and/or sufficient conditions, or a cluster of conditions, Ci....Cm, and that at a later point, in a competing theory T', concept "X" is "defined" by Ci....Cn or Ci....Ck. How are the "X's" related? Are they related at all? What can we say about it? According to the traditional conception presented above, we would have to say for example that we are dealing with two different concepts in two different theoretical contexts, that the meaning of "X" has radically changed, or else, that ["X" in T'] has replaced ["X" in T]. If taken in this sense, the traditional conception is decidedly ahistorical, and as such, it contributes to the problem of incommensurability. However, one may argue that it is, after all, possible to "compare" the developed concepts ["X" in T] and ["X" in T'] by specifying some degree of similarity of meaning, some degree of overlap in the conditions "defining" them respectively. This might give us some basis for commensurability and comparison of "developed and finished products." But even in this case, we would still have no way of telling about the process which led from one product to the other, that is, of characterizing the transitional development(s). This becomes all the more problematic when we are dealing with
developed concepts (e.g., "mass") from two radically different scientific theories (such as Newtonian and Einsteinian mechanics) [See Part III: Shapere, Nersessian, and Kuhn, 1988 for a similar point].

Though developed concepts from competing scientific theories may be different, one would want to be able to account for the sometimes long and protracted process by which scientists have gone from one to the other. In other words, one would want to be able to say how much of their respective meaning they share with their respective predecessors -- with the understanding that each concept will share more with its immediate predecessor than with those more remote. Thus rather comparing two developed concepts (e.g., "mass") in two different theories (e.g., Newton’s and Einstein’s), we need to be concerned with all the transitional developments which have taken place in-between. To take another example, one would want to be able to show for example that Einstein’s concept of "electromagnetic field" shares more of its meaning with Lorentz’s corresponding concept (or Maxwell’s) than it does with Faraday’s -- and not just "compare" the first and the last one (Shapere, 1984, 1989; Nersessian, 1984, 1987). More generally, one would want to be able to show that while there is meaning-variance and discontinuity in the process of meaning-making in science, there is also a significant
(and perhaps greater) degree of commensurability and continuity. A necessary component of an adequate account of the process of meaning-making in science is an adequate conception of meaning, in other words, an adequate concept-representation. 6

Since, as we have seen, the traditional conception of meaning and concept-representation is inadequate on both of the points mentioned above, the question then is whether a cluster theory of meaning and cluster concept-representation, like the ones proposed by Gasking, Putnam, or Achinstein, fares better. Or whether it is necessary to formulate a new and revised version of cluster theory.

Before turning to an examination of these traditional versions of cluster theory, it is only appropriate to set up the stage by giving a general reconstruction of the traditional version of cluster theory and pointing out the main objections commonly raised against it. 7 Tentatively, this should also enable me to sketch out some of the features of the adequate cluster theory as I see it (see Part IV).

According to the traditional version of cluster theory, the meaning of a term is determined by the cluster of properties associated with it, enough of which must be satisfied for the term to be applicable. A speaker or
individual member of a linguistic or scientific community is said to understand the meaning of the term and to know how to use it, if s/he knows the cluster of properties associated with it, and furthermore if s/he knows "enough" of the properties which must be satisfied for the term to be applicable. This is obviously very reminiscent of the conditions of application of the early "verification theories of meaning"—except that the cluster device, and not that of set, is introduced.

Given such a formulation of cluster theory, at least the following objections can be raised against it: (i) It assumes the cluster of properties associated with a term determines or constitutes its meaning, and thus appears to be still committed to some form or grade of essentialism. (ii) It is vague concerning the number of properties required in the cluster; "enough" does not say how many is sufficient. Related to this problem is another one, namely, how far wrong can the cluster—composition be and yet still refer to whatever the corresponding term was intended to pick out? (iii) It assumes that the locus of meaning (and understanding) is the individual. (iv) It assumes that the individual must have some sort of epistemic access, i.e., know the cluster of properties associated with a term in order to be able to use it successfully.8

By way of transition, it might be helpful to summarize
the mains objections which have raised against traditional theories:

(1) Inadequate Concept-Representation/Conception of Meaning.


(3) Infectious Semantic Holism/Essentialism of Meaning Paradoxes of Incommensurability.

In anticipation, I will show (in Part IV) that the new cluster theory does not make any of the assumptions of traditional theories of meaning which justify these objections against them. In the meantime, I will turn next to various traditional versions of cluster theory. I will show that though they each make some improvement upon traditional contextual theories of meaning --by virtue of drawing some of the implications of the "cluster" device, they each fail to provide an adequate account of the process of meaning-making in science.
Chapter 4: TRADITIONAL VERSIONS OF CLUSTER THEORY

It should be noted at the outset that these versions of cluster theory, those of Gasking [1960], Putnam [1962], and Achinstein [1968], shall be reconstructed here with my present purpose in mind, namely to determine their strengths and weaknesses as possible accounts of the process of meaning-making in science. It might be helpful to note that each of these versions offers a more or less detailed analysis of the structure (logical, ontological and semantical) of a "cluster." Each claims that the structure of a cluster is determined or rather constrained by a fundamental principle. In my reconstructions, I will bring out and formulate as clearly as possible the principle on which each of these versions of cluster theory is presumably based. Ultimately, my goal remains however to evaluate them against the background of the objections raised earlier.

A. Gasking's Version of Cluster Theory

Gasking's notion of "cluster is an intricate and difficult one. At first, one may be tempted to think of it in terms of sets or classes. But, as he points out (1960, p.1) his concern is not so much with sets (i.e., aggregates in the extensional sense, comprising timeless entities, timelessly interrelated) or with classes (i.e., aggregates in the intensional sense, defined by an attribute or set of attributes) as it is with referent classes, that is, classes with variable membership, defined by symmetrical and
transitive relations and a focus. A referent class consists then of all those entities which have the defining relation \( R \) to the defining focus (whichever it is), and the focus itself counts as a member of the referent class (1960, p.6). 2

Now, one may ask: but what is exactly a cluster? What is its structure? As a preliminary toward a general definition, it may be appropriate to characterize it as follows:

(I) A cluster is a set of entities \( S_n \) (e.g., attributes, properties, states, events, forms, episodes, objects, activities, etc) which is, at some time \( t_i \), integrated by a symmetrical and transitive relation \( R \), and forms thus an omnifocal set with respect to that relation. 3

By way of further clarification of the logic of clusters, we should also keep in mind that:

(a) A set of entities \( S_n \) is said to be integrated by a relation at some time \( t_i \) when all of its members have \( R \) to each other, and none has \( R \) to any non-member.

(b) A set of entities \( S_n \) is said to be omnifocal with respect to a relation \( R \) when to each of its members there corresponds a definition, in terms of \( R \) and that member as focus, which designates the entire set \( S_n \).

Given (a) and (b) as defined, it follows that

(c) A set of entities \( S_n \) is omnifocal with respect to a relation \( R \) iff it is integrated by that relation \( R \) (1960, pp.6-7).

One can thus use either integration or omnifocality as a characterization of a cluster, and subsequently think of it either as an integrated set or as an omnifocal set.
It is my understanding that Gasking thinks of it in the latter way (1960, p.12). Indeed, his conception of cluster is one which is based on what I shall call the principle of omnifocality. To think of a cluster as an omnifocal set as depicted above has, however, some important implication. These can be formulated as follows: (i) A cluster has as many "definitions" or rather "descriptions" as it has members, since it is definable or describable in terms of any member. (ii) Any such definition or description may be regarded as just as correct as any other, provided it always has, for its focus, some member of the cluster. This may be, but it is hard to believe that they are equally useful or satisfactory. Also, it seems to be questionable since it would dissolve any meaningfulness to the distinction between central and secondary properties. (iii) No one member of cluster (or omnifocal set) need be essential (i.e., necessary and/or sufficient). This claim constitutes one of the original motivations for introducing the notion of "cluster" in the first place. (iv) Assuming that we have a stable state of affairs, that is without any change in membership, the members of a cluster at a given time can be identified by anything that is then a member, provided one knows the defining or integrating relation. This is I presume a consequence of Gasking's principle of omnifocality. (v) For the purposes of human interactions and communications, all this means that there is no need for
everyone to agree on one criterion (offered by the definition, specifying necessary and/or sufficient conditions) as the only "correct" one to follow in applying a cluster-name, and also no need for any given person always to stick to the same criterion. Instead, there is a large number of possible criteria (depending on the number of members in the cluster-set) which different people, or even the same person, could adopt, at different times, all such criteria (and any of them) being in principle equally good for picking out the members of the cluster-set. This seems to be quite plausible and defensible, even if one rejects some of the other claims made above.

So far we have concerned ourselves only with a "stable state of affairs" and thus have given a tentative characterization of a cluster, i.e., (I) above. But this does not yet constitute a full definition, for, as Gasking puts it, "until we have considered the ways in which a cluster can change while yet remaining the 'same cluster,' this notion will not be fully explained." (1960, p.12).

In order to fully explain the notion of cluster, he proposes the following model which involves a changing state of affairs:

Suppose that twenty-six ships, whose names A, B, C, D, etc., are at sea in war time maintaining radio silence. Each is in visual contact, by means of signal lamps, with at least one of the others, so that every one of the twenty-six is serially in contact with the remaining twenty-five, but with no other ship. If that situation
were to last for any appreciable time we should think of the twenty-six ships as constituting a single enduring entity—a cluster of ships. Since we might want to refer to this cluster on several occasions we might perhaps give it a proper name, such as 'Tom'. (Remember that weather-forecasters sometimes give such names like 'Clara' to hurricanes, cyclones and typhoons).

If a storm scattered our twenty-six ships, so that none was any longer in contact with any of the others, we should have to say that Tom had ceased to exist. [But] suppose it came to pass that thirteen ships A to M were serially in contact with each other, and that the ships N to Z were also serially in contact, but that no ship of the first set was in contact with any member of the second. In that case, too, we should have to say that Tom was no more—that it had split into two new clusters (say 'Tim' and 'Tam'), neither of which could properly claim to be the original Tom. If the split was only temporary, and the whole twenty-six were very soon again in serial contact, we might well say: 'For a short while Tom split into two, but now it's back to normal again' (Gasking, 1960, pp.14-5, my addition in last parentheses).

What does this model tell us about clusters? First of all, we learn that for the notion of an individual cluster (to which we could give a name) to get a grip, to coalesce meaningfully, a certain minimum degree of stability is required, because too frequent and too complicated fissions and fusions would be defeating. But, as Gasking hastens to point out, "the stability (here in question) need not be absolute" (p.15). He goes on to explain what he means in a manner which shows that a cluster can change in membership over time and yet still remain the 'same cluster'. The main point is that even if Tom were to lose or gain a few members (one, two, perhaps three and even more—it is not clear how far one can go), we would still think of it as the cluster Tom. He writes in this respect:
If A were to drift away from the other twenty-five ships, while these latter still remained serially in contact with each other, we should no doubt think of the cluster, Tom, as still existing. We should say that Tom had lost one of the ships it used to consist of, and now consisted only of the twenty-five ships B to Z. And if, later on, these twenty-five were joined by two new ships, * and #, we should say that Tom had grown by the addition of two new ships, and now consisted of the twenty-seven ships * and # and B, C, D, E, etc. If one of these twenty-seven—the ship M, say—were to drift away, leaving the other twenty-six serially in contact with each, we should say that Tom was still there, but that M no longer belonged to it (1960, p.15, Greek letters alpha and beta in text instead of * and #).

Secondly, as the passage just quoted indicates, we learn that, of the twenty-six ships, A to Z, of which Tom at present consists, no one is essentially a member of Tom. Thirdly, insofar as what holds for actual changes in the membership of a cluster holds as well for changes in our beliefs about the membership of the cluster in question, we conclude that no ship which is, or thought to be a member of Tom, is essentially a member; and no ship which is, or thought to be a non-member, is essentially a non-member. We can see how this is so by attending to Gasking’s reasoning in this instance. He writes:

Suppose that you believe that the ships A, B, C, D, E, and F, with some twenty or so others, constitute a set which is integrated by the relation serially in contact with, and that you call this cluster by the name ‘Tom’. Now suppose that you discover that your sources of information were defective—that the ship A is not, as you had thought, serially in contact with any of the ships B, C, D, E, or F, but that these are all, along with twenty or so others, serially in contact with each other. In such case you would still give the name ‘Tom’ to the cluster consisting of B, C, D, E, F and the rest. And you would say: ‘I used to think, mistakenly, that A
was a member of Tom, but I have since discovered that it is not.' Likewise, you might discover that *, which you had thought to be far away, had all along been in contact with B, C, D, E, F and the rest. You could say, if so: 'I did not realize that * was a member of Tom, but now I have discovered that it is.' And what applies to A or to * applies to any other supposed member or non-member (1960, pp.15-6, italics in text).

From the model which I have been discussing, we also learn about two of the most important features of the logic of clusters. These are: (i) open-texturedness: one or two (perhaps more) elements of a cluster may be added or left out without dissolving the existence of the cluster in question, or the 'meaning' of its corresponding cluster-name. (ii) Corrigibility: one might be led to correct or revise one's conception of a cluster by learning new facts about its composition and membership, -- or else, as Gaskin needs to add, by deferring to or by studying other people's usage (experts, relevant authorities). By "conception of a cluster," I mean simply the "web of beliefs" that one may have about its composition or membership.

Given all the above, we can now reformulate the definition of a cluster given earlier (i.e., (I)), by taking into account change over time, in the following way:

(II) A cluster is a sequence of sets of entities Sn = {{Si,ti},........{Sw,tw}}, such that most of Sn is, over an appreciable period of time, integrated by a relation R (transitive and symmetrical), and forms thus an omnifocal sequence of sets with respect to that relation.
Though Gasking nowhere gives such a definition explicitly, I think it captures his characterization of a cluster over time.

As one may have guessed, all that has been said so far about clusters and cluster-names in general can be readily applied to the representation of concepts, including scientific concepts. Thus, one can define a cluster representation of concepts which is both synchronic and diachronic in the following manner:

(I*) A cluster-concept is a set of properties or attributes Sn which is, at some point in time ti, integrated by a relation R (transitive and symmetrical), and which forms thus an omnifocal set with respect to that relation.

(II*) A cluster-concept is a sequence of sets of properties or attributes Sn such that most of Sn is, over an appreciable period of time, integrated by R, and forms thus an omnifocal sequence of sets with respect to that relation.

On the basis of these definitions and the features of clusters that I have presented throughout, we can also characterize how one identifies the members of a cluster-concept and also cluster-concept membership. To identify the members of a cluster-concept, all one needs to know is (i) the integrating relation and (ii) one (any) present attribute of the cluster-concept. As for cluster-concept membership, one says that x is a member of a given cluster-concept if and only if x has most of (the set(s) of) attributes integrated by a relation R (typically associated with it) over an appreciable period of time.
Having characterized cluster concepts rather abstractly so far, it is now appropriate to consider a few examples. But first let us note that though not all concepts are cluster concepts, a good number of them seem to be, at least if we believe Gasking on this matter (1960, pp.25ff). In his view, concepts having to do with material bodies (e.g., houses, knives, human bodies), species (both morphological and biological), states and properties (e.g., diseases), objective events, cultural kinds (e.g., games), and chemical natural kinds (e.g., gold, iron, copper, sulphur, etc) can all be construed as cluster-concepts. For Gasking however, there are different kinds of cluster concepts (e.g., "simple clusters," "chain clusters," "essential and accidental clusters," 1960, pp.9-10;13-4). The latter distinction seems problematic to me since the original motivation for introducing clusters in the first place was to get away from any notion of essentialism. If, in addition, we believe Brian Ellis (1969, pp.34-5), then we would have to hold that our quantity concepts in science (e.g., "length," "kinetic energy," "velocity," "density," "electromagnetic field") are also cluster-concepts. (See Nersessian, 1984,p.158; Achinstein,1968, chap.1 and 2 for further discussion of cluster-concepts).

As we have seen, my reconstruction of Gasking's version of cluster theory shows that there are some significant
differences between the traditional concept-representation and the cluster representation that he has in mind. While the former represents a concept at some point in time as a set of necessary and/or sufficient conditions, the latter represents it as an omnifocal set of attributes (which are neither necessary nor sufficient, singly or jointly); as a consequence, none of the implications mentioned with respect to the latter holds with respect to the former. While the former completely overlooks the development over time, the latter cannot be made sense of without it; as a consequence, the features of clusters, cluster-names and cluster-concepts, for that matter, do not apply to concepts as traditionally represented. In this sense, the cluster representation seems to be closer to capturing the open-endedness of concept-formation and of the process of meaning-making in science. This constitutes a significant factor in its favor.

But is this to say that Gasking's theory is satisfactory? I do not think so, for at least one major reason, and that is, it is based on the principle of omnifocality as I have tried to show in this discussion. It undoubtedly makes an important point, namely that concepts (including some scientific concepts) need not (and perhaps cannot) be defined in terms of a set of necessary and/or sufficient attributes, but instead in terms of a cluster of
correlated attributes (which includes the dimension over time). However, to require that a cluster of attributes forms an omnifocal set is to reintroduce the "essentialism of meaning" that characterizes the traditional concept-representation. Furthermore, it is questionable whether these attributes are, so to speak, on a par, have the same value or weight as far as the identification of the instances of a cluster-concept is concerned. The whole point of bringing in the "cluster" device was to eliminate any trace of essentialism and introduce some flexibility in the structure of concepts, particularly over time. As I will argue in Part IV, all that Gasking needs to do is to adopt instead a more reasonable principle to characterize the internal structure of clusters: what I shall call the principle of correlation. Briefly, according to such a principle, the attributes forming a given cluster are co-related in some fashion or other, which is to be determine for each concept. They "cluster" together by virtue of the relation(s) holding between them.
B. Putnam's Version of Cluster Theory

Putnam's discussion of cluster-concepts (1962, pp.50-4) is sketchy or rather "programmatic," as some might want to say. It deals specifically with a scientific example. In this section, I propose to examine briefly some of the implications for an account of meaning and meaning-change in scientific theories of a statement that he makes, according to which "most of the terms in highly developed science are law-cluster concepts" (1962, p.52, my underlining). But first, it is important to have some idea of the context in which his discussion of such concepts arises.

According to Putnam, the notion of cluster concept has been introduced by philosophers (of language) in order to deal with a particular kind of problem. This kind of problem comes about because of the traditional representation of a concept as a set of attributes. Take the concept 'man' and suppose that it is defined by a list of attributes Pi....Pm, where Pi= "is rational" and Pj= "does not grow feathers all over his body." One can ask quite legitimately "can there be a man without Pi, or Pj, and so on?" That is, can there be a man who is irrational from birth and/or who grows feathers all over his body? In each case, the answer is almost certainly "yes," and yet, according to the traditional concept-representation, we may be expected in some cases to say that we are dealing in each case with a different concept, or else that the meaning of
the concept has radically changed. But, as we can easily see, this is rather counter-intuitive.

In order to resolve this sort of difficulty, the notion of cluster concept was introduced. The meaning of a concept is no longer given by a set, but by a cluster of attributes. Thus, the concept "game" is no longer defined in terms of a conjunctive set of attributes such as following rules, being played, having a winner and a loser, being entertaining, involving the scoring of points, and so on, because some perfectly acceptable games lack some of these attributes. Instead, something is a game (or a man) if it has many, most, or in any case enough of the attributes of a cluster. If a large number of these attributes were to be left out such that the extension (or reference) of the concept were to be radically changed, then this would constitute an arbitrary change in the meaning of the concept. But, according to Putnam,

if most of the attributes in the cluster are present in any single case, then under suitable circumstances, we should (be inclined to) say that what we had to deal with was a man [or a game, depending on the case being considered] (1962, p.52; my addition in parentheses).

Generally speaking then, it appears that a cluster concept in Putnam's view, is a concept whose meaning is given by a cluster of properties, not all of which must be possessed by an individual or entity in order to fall under the concept's reference or extension, although it must
possess a **good number of them** (perhaps most, it is not clear).

This characterization is reminiscent of Gasking's in a number of respects. For example, it follows from it that a cluster concept is open-textured; one or a few of the attributes or properties that serve to determine or constitute its meaning may be left out or added without involving a change in its meaning. Moreover, since satisfying a certain proportion (most or a good number) of the attributes of a cluster is a sufficient condition for its being applied to an entity, a cluster concept is associated with several possible "descriptions" for determining its extension or reference. No one of these descriptions is essential, but all are more or less serviceable (depending on the context and the case) for picking out what goes into the extension or reference of the concept.

In analogy with the notion of cluster concept, Putnam introduces what he calls a **law-cluster**. He writes:

Law-cluster concepts are constituted not by a bundle of properties as are the typical general names like 'man' and 'crow' (and 'game'), but by a cluster of laws, which, as it were, determine the identity of the concept (my underlining).

And he adds, significantly enough, from the point of view of an account of meaning and meaning-change in scientific theories:
The concept of (kinetic) 'energy' is an excellent example of a law-cluster concept. It enters a great many laws. It plays a great many roles, and these laws and inference roles constitute its meaning collectively, not individually. [I want to suggest that most of the terms in highly developed science are law-cluster concepts] (1962, p.52; my underlying).

It is not quite clear what Putnam means to say by "these laws and inference roles constitute its meaning collectively." Does he mean that they all constitute its meaning or just that most constitute its meaning collectively but not conjunctively, and perhaps rather disjunctively? In virtue of the analogy between a cluster concept and a law-cluster concept, it might be more charitable to interpret Putnam as meaning the latter.

Subsequently, since an individual or entity need not possess all of the attributes of the cluster concept in order to fall under that concept, any one of the laws which (collectively, with the others) constitute the meaning of the concept of 'kinetic energy' may be abandoned without destroying its identity. Presumably, this is a significant characteristic of law-cluster concepts, given Putnam's insistence on it. It can be formulated more generally as follows: one or a few of the laws which constitute the (intensional) meaning of the concept may be abandoned or added, for that matter, without altering the (extensional) meaning or identity of the concept.

It is important to note here that Putnam (1962), contrary to the tradition which construed meaning primarily
as sense or intension, construes meaning primarily as reference-fixer or extension-determiner. Thus, he able to say that minor changes in the 'meaning' need not change what it does, that is, fixing reference. Perhaps it would be more accurate to say, in light of his later proposals (1973, 1975; see Part II), that he urges us instead to turn to 'reference' in order to accomplish the work formerly accomplished by "meaning." In applying his analysis to 'kinetic energy', he concludes that the 'meaning' of the concept has not changed with Einstein's theory of relativity, because its extension has not changed. His argument runs roughly like this: If the extension had changed, then the extension of the concept of 'energy' would have changed as well. Kinetic energy is only one kind of energy, which can be transformed into other kinds and vice versa; subsequently, an adequate physical theory cannot change the meaning of the concept 'kinetic energy' without at the same time changing the meaning of the concept 'energy', that is, without giving up altogether the idea that 'kinetic energy' is literally one kind of energy. But the extension of the concept 'energy' has not changed. The forms of energy and their behavior are the same as they always were, and they are what physicists talked about before and after Einstein. Thus, the meaning of the concept of 'kinetic energy' has not changed. (Putnam, 1962,p.52-3).
I find his argument rather unconvincing. In particular, I think that the crucial assertion starting with "subsequently..." is questionable: does a change in the meaning of "kinetic energy" entail a change in the meaning of 'energy,' i.e., that we must give up altogether the idea that 'kinetic energy' is literally one kind of 'energy'?

In spite of this however, I feel that Putnam’s argument can be defended in some way. Even though it is not quite appropriate to talk about the intension of a law-cluster concept, one could say that the kinetic energy laws introduced by Einstein's theory of relativity did alter the intension of the concept of 'kinetic energy,' but it did not alter the extension of the concept. The extension of 'kinetic energy' was literally the same after the change just as it was before: the kinetic energy of a particle remained the energy due to its motion. Insofar as meaning is construed extensionally, one can conclude that the meaning of the concept 'kinetic energy' has not changed.

For those familiar with both classical and relativistic mechanics, the same point is, I think, made in a more technical fashion, by Achinstein (1968, pp.55-6, my underlining and parentheses) as follows:

In many presentations of classical mechanics, the definition of 'kinetic energy' as \(1/2mv^2\) is arrived at by considering the work done by a force on a particle during the displacement from an initial position \(P_0\) to a final position \(P_1\). Work is defined in vector notation as
\[ W = \int_{P_0}^{P_f} F \cdot dr, \text{ where } F \text{ is force and } dr \text{ is an} \]
\[ \int_{P_0}^{P_f} F \cdot dr = 1/2 m v_i^2 - 1/2 m v_o^2. \]
This means that the
work done by the particle during the displacement from \( P_0 \) to \( P_f \) is equal to the difference between the value of the quantity \( 1/2 m v^2 \) at the end of the displacement and its value at the beginning, where \( m \) is the mass of the particle, \( v_o \) its initial velocity, \( v_i \) its final velocity. A
knowledge of this derivation, or some other like it, (as we
shall see in relativistic mechanics) is crucial for
understanding how the expression \( 1/2 m v^2 \) is obtained. 
Otherwise, the choice of this expression would seem
perfectly arbitrary. Furthermore, it is also crucial for
understanding what has changed in the definition. In
relativity theory, the 'kinetic energy' (of a particle of
rest mass \( m_0 \) and velocity \( v \)) is defined as
\[ \frac{1}{m c^2} \left( \frac{1}{1 - \frac{v^2}{c^2}} - 1 \right), \text{ where } c \text{ is the velocity of light.} \]
This definition is also obtained by considering the work
done by a particle during the displacement from \( P_0 \) to \( P_f \), that is,
\[ \int_{P_0}^{P_f} F \cdot dr. \]
But in relativistic mechanics, \( F = \frac{d}{dt} \left[ \frac{m v_o}{\sqrt{1 - \frac{v^2}{c^2}}} \right] \).

As I understand this passage, it seems to further
illustrate the idea that a law (or set of laws) taken to be
"definitional or stipulative in character" of "kinetic
energy" in classical mechanics is abandoned and replaced by
another in relativistic mechanics and yet, the identity (or
extensional meaning) of the concept has, in a certain
respect, remained. Interestingly, I find this expressed
also in a physics textbook (see Tipler, 1976, pp.176ff; 684ff, for a simpler account of the derivations mentioned by
Achinstein above), in which a discussion of relativistic
energy begins as follows: "As in classical mechanics, we shall define kinetic energy as the work done by an unbalanced force in accelerating a particle from rest to some velocity" (p.686, my underlining). Thus, in both classical and relativistic mechanics, the kinetic energy of a particle is the energy due to its displacement from an initial position Po to a final position Pi; what has changed from one to the other is (just) the way of characterizing, calculating, and deriving it. That is to say, once again, what has changed is the 'intension' while the extension has not.

At least one other objection may be raised against the view that Putnam suggests. One may argue that it fails to specify a sufficient condition for meaning-change, i.e., a criterion for determining at what point there is meaning-change. This would be however to mis-understand the whole point about construing meaning as reference-fixer or extension-determiner (Scheffler 1967, pp.54-66 [Part II]; see also Quine, 1953, pp. 47-8; 132). Furthermore, it is important to keep in mind that Putnam's account is based on the principle of over-determination: a law-cluster concept (e.g., 'kinetic energy') is such that the laws which constitute its meaning or identity do not merely determine its extension, they also overdetermine it in several ways. As a result, the deletion of any one or a few of the laws constituting its meaning or identity does not necessarily
alter its extension. And so, it somehow makes sense to talk about meaning-change (in the intension) without meaning-change (in the extension). Now does this mean conversely that the cluster of properties or laws whatever the case may be is under-determined by the reference or extension, and that, as a result, it is possible to have cases where (slightly) different clusters of laws or properties can still pick out the same referent or extension? It seems to me that this must follow. Furthermore, one can argue on Putnam's behalf that in fact a specification of a sufficient condition for (extensional) meaning-change can be obtained. The laws constituting the meaning or identity of a law-cluster concept do not just (over)determine its extension, they also assert various factual connections purportedly holding between entities falling under the concept and other entities. These assertions about factual connections are not exactly part of the identity or extensional meaning of the concept, but they affect and are affected by the extension of the concept. Thus, the addition or removal of laws from the cluster will alter the extension only if the factual connections require alteration in the concept's extension. This would be the case, for example, if the new law(s) asserted factual connections which could not be true given the previous extension of the concept.
A major implication of viewing scientific concepts as law-cluster concepts should by now be easily discernible. The (extensional) meaning of a law-cluster concept is not altered by the addition or deletion of any one or a few of the laws which constitute it—except as specified above. Such change is not always meaning-change, contrary to what one might hold in this respect; that is to say, it cannot always be explained as "replacement," as "definition" of a new concept, or else as a radical change in the meaning of the concept. Admittedly, Putnam recognizes that there is a sense of meaning in which "one might say that the change in the status of a law or principle brings about a change in the meaning of the concept," but, as he hastens to add, "this must be a rather fuzzy sense of meaning". He notes:

This 'fuzziness' is evidenced by the fact that although one can say that 'kinetic energy' has a new meaning, one cannot say that 'kinetic' has a new meaning, or that 'energy' has a new meaning, or that 'kinetic energy' is a new idiom (1962, p.53).

What one should say rather, in Putnam's view, is that the meaning of the concept has not changed enough to affect 'what is talked about' i.e., the reference or extension of the concept. And one should expect that an adequate theory will take this into account.

Before closing this discussion, one should ask whether a more fully developed cluster theory based on a principle of overdetermination, like Putnam's, does or can provide us with an adequate account of the process of meaning-making in
science. If what Putnam suggested is right, namely that "most of the terms of highly developed science are law-cluster concepts," including the principles of geometry which he discussed along with the "definition of kinetic energy," then such an account would have a rather significant scope. And this is not negligible when it comes to evaluate competing accounts, as we all know. Also, as we have seen, it attempts to defend the following fundamental idea, namely, that the concepts of a new theory do not always completely replace those of a previous one. As a result, unlike the traditional view which ends up with extreme incommensurability, it does not equate meaning-change with replacement of one theory with another one, incommensurable. By construing meaning extensionally and by characterizing it as overdetermined in several ways by a cluster of laws or properties, it seems to favor an approach which would account for a significant degree of commensurability by providing for the continuity of extension or reference while admitting only mild forms of "meaning-variance" or what may be called "incommensurability of intension or sense." [This seems to announce Putnam's later views, 1973, 1975]. But is not this simply Frege's idea, couched in different terms, i.e., that different senses (or "meanings") may still pick out the same reference [while a reference-change presupposes different senses]?
In the end whether or not a cluster theory, like Putnam's, provides an adequate account of the process of meaning-making in science remains to be seen. To pronounce a final judgement, we have to wait until Achinstein's version is examined next. For, as Suppe remarks quite appropriately, Achinstein presents "an extremely lucid and careful formulation of cluster concepts, which can be construed as including law-cluster concepts," and thus as providing "a fuller characterization of Putnam's key notion." (Suppe 1977, p.74, note 154; see also 1973, 1989 for application and extension of Achinstein's version to the problem of taxonomy and its implications for philosophy of science). In the meantime however, I should stress that any adequate theory or account must incorporate the main insight of Putnam's account.
C. Achinstein's Version of Cluster Theory

I now turn to Achinstein's account. Even though Achinstein nowhere uses the term "cluster," it is clear from the content of his discussion that he attempts to develop a version of cluster theory. I will first articulate the main features of this theory and secondly, evaluate it, after considering a number of possible objections.

Achinstein's theory in Concepts of Science (1968) is concerned primarily with "concepts important for understanding the nature of science" (p.vii), namely definitions, theories, and models. The particular context in which it emerges is the discussion of definitions of scientific terms (1968, pp.1-66). By "definition" Achinstein means to say "the sort of information contained in a scientific dictionary entry," 2 (1968, p.1) typically listing various "properties" 3 about a given item which is referred to by a term. For example, for the term "copper," we may have the following sort of information: a reddish metal of atomic number 29, with a melting point of 1083c, malleable, ductile, a good conductor of electricity and heat, with a density of 8.92g/cc, found in abundance in its natural state, used for wiring, coins, piping, pottery, has been used since ancient times, etc.

The main question that Achinstein addresses is this: what is the relationship between the information presented (i.e., the properties cited) in the definition of a
scientific term and the term defined? He believes that an answer to this question will permit a better understanding of the nature of scientific definition(s). But might this in turn lead to a better understanding of the nature of the process of meaning-making in science?

Despite the variety of scientific terms and the fact that a definition of a scientific term may vary considerably from dictionary to dictionary, Achinstein sets about to show that some general claims about such definitions can be made.

For heuristic and explanatory purposes, Achinstein distinguishes three kinds of scientific terms, namely A-terms, B-terms, and C-terms. But he adds: "These are not meant to exhaust the types of terms in science; nor is the division perfectly sharp" (1968, p.2).

The A-terms designate physical objects, phenomena, or stuffs, e.g., copper, gold, metal, electron, atom, acid, oxygen, electroscope, insect, molecule, etc.

The B-terms express somewhat more abstract concepts applicable to physical objects, stuffs, phenomena and so forth, provided the latter satisfy certain necessary and sufficient conditions. They include for example conservative system, diatomic molecule, Newtonian system, reversible process, Carnot cycle, rigid body, Bohr atom, quadruped, fermion, etc.

The C-terms designate quantities capable of numerical degree, e.g., kinetic energy, velocity, length, temperature, entropy, density, specific gravity, force, potential energy, etc (see 1968, p.2).

Achinstein's main goal is to introduce the notion of semantical relevance, and to show how, in many important
cases, this concept instead of the traditional ones of logical necessity and sufficiency must be invoked to understand the nature of definition(s) in science. More specifically, with respect to the different kinds of scientific terms, he proposes to show that, while B-terms admit of explicit nominal definitions in terms of logically necessary and sufficient conditions, A-terms often admit instead of what I shall call "cluster linguistic descriptions" in terms of relevance/semantical relevance. As for C-terms, some are "defined" similarly to A-terms, while others are "defined" similarly to B-terms.

Having shown that the relationship between a property of an item designated by an A-term and being an item of that sort is not, in general, one of logical necessity and/or sufficiency, Achinstein takes up the next question, namely how can it be described? For this purpose, he introduces first the notion of relevance. Relevant properties for a term are those properties such that having one of them counts, to some degree, for or against the term applying to an object (1968, p.6). On the basis of this notion, Achinstein draws a distinction between semantical and nonsemantical relevance as follows:

Suppose Pi,...,Pn constitutes some set of relevant properties of X. If the properties in this set tend to count in and of themselves to some extent, toward an item's being classifiable as an X, I shall speak of them as semantically relevant for X. If the possession of properties by an item tends to count toward an X-
classification solely because it allows one to infer that
the item possesses properties of the former sort, I shall
speak of such properties as nonsemantically relevant
1968, pp.8–9, italics in text; my underlining).

In other words, semantically relevant properties are those
properties which are relevant in and of themselves. But
what exactly does Achinstein mean by "properties relevant in
and of themselves"? I will come back in a moment to this
question.

To illustrate in the meantime what Achinstein has in
mind, let us consider the term "copper." Given my earlier
characterization, "having a melting point of 1083°C" is
certainly relevant for copper, but it is not logically
necessary or sufficient. The fact that an item with certain
other properties has this melting point will normally count
as some reason in favor of concluding that it is copper
(without necessarily settling the matter); and if the item
has sufficiently many (most) other properties cited in the
dictionary, the fact that it melts at 1083°C might
justifiably be held to settle the question of its
classification. Other relevant, though not logically
necessary or sufficient, properties would be reddish color,
having a density, being a good conductor, malleable,
ductile, etc.

But what are we to say about "having the atomic number
29"? Is it another relevant or rather semantically relevant
property, in Achinstein's sense —though not logically
necessary and sufficient? In anticipation, I think that Achinstein' view on this point is not very clearly spelled out. In any case, this is a matter open to debate. Some scientists and philosophers of science would want to claim that, given the prevailing and established scientific knowledge about atomic theory, the property here in question is to be considered as logically necessary and sufficient, i.e., as essential. Some even go on to argue that this kind of essentialism is not unscientific (see: Leplin, 1988).5 However, others, such as Putnam (1962) and Shapere (1984) might disagree. They might argue that reasonable hypothetical cases and scenarios could (and should) be envisaged (involving more or less radical modifications of atomic theory) in which the following obtains. This would accord the open-ended nature of the scientific process, and would confirm what we have learned from the history of science. Firmly entrenched scientific beliefs and claims can be rejected as many have already been rejected in the past. First, one may discover a substance sufficiently like copper in other respects but with a slightly different "atomic number," which we nevertheless classify as copper. Second, we may discover a substance possessing the atomic number 29 but lacking a large number of other properties cited in the definition of "copper," which we do not classify as copper. In the first case, having the atomic
number 29 would not be logically necessary. In the second case, it would not be logically sufficient. It is therefore an important adequacy requirement for an account to be in accord with the open-ended nature of scientific inquiry, and with what we have learned from the history of science.

Before proceeding any further, two questions must be addressed: First, how can we determine whether properties are of one sort or the other (semantically/non-semantically relevant)? Second, what sorts of properties will be semantically relevant for items designated by A-terms?

Departing from positivist strictures, and in a way closer to the historicist tradition, Achinstein's answer to the first question is that "we must look to what scientists do in actual situations of classification, to what, in fact, they say about such situations; we must also determine what they would do and say about hypothetical ones" (1968, p. 8). This clearly indicates that he wants his notion of semantical relevance to be ultimately "anchored" in the scientists's practices and procedures. But as I will argue, Achinstein seems to undermine his own wish.

Given Achinstein's characterization of semantical relevance in terms of relevant properties, which are relevant in and of themselves (1968, pp. 8-9), one might argue that, in his view, semantical relevance requires irreducibility, or else, that reducibility rules out
What makes a property relevant "in and of itself"? As noted earlier, a relevant property gives us information about whether or not a term applies to object which have the property. That information may be relevant because it gives information about other relevant properties. If so, the original property is relevant not in and of itself, but rather through other properties. For example, "mined in South Africa" is relevant for "copper," but it is not relevant in and of itself. Similarly, and in light of previous remarks, "has the atomic number 29" is relevant for "copper," even centrally relevant --given currently established scientific theory, but is it relevant in and of itself, i.e., semantically relevant?

Let us suppose for the sake of argument that a property P relevant for term X is irreducible (i.e., semantically relevant or relevant in and of itself) iff there are no other properties Pn relevant for X such that the reason P is relevant to X is that P gives us information about Pn. Suppose further that in the cases where we can specify the properties Pn in terms of which P is reducible, we shall say that P is reduced. The possibility that properties are reducible but not reduced is more than academic and should be stressed. Since furthermore reducibility may come into play in at least two ways for the kind of A-terms here in
question, Achinstein's view faces two major problems. (i) The fact that we may succeed in reducing specific properties does not imply (the belief) that other properties are themselves irreducible. Future scientific inquiry may yet reveal more fundamental properties. In the scientific study of natural objects and phenomena, does it make sense to ascribe absolute irreducibility to any specific property? Certainly not, if we consider seriously the open-ended nature of scientific inquiry. (ii) Suppose however that we discover specific properties that were irreducible parts of some natural object or phenomenon, the reason that we would associate these properties with the term would still be that these properties are parts of the natural object or phenomenon in question. In light of (i) and (ii), we must conclude that properties associated with "cluster terms" are reducible on both epistemological and semantical grounds.

The question then is: can a property which is reducible, i.e., has reasons for its relevance, be relevant "in and of itself", i.e., semantically relevant? If irreducibility is required for semantical relevance, then terms like "copper," "metal" and "oxygen," as used, do not have properties which are relevant in and of themselves. Just as it is not part of our conceptions of these terms that the properties associated with them are logically necessary or sufficient, it is not part of our conception that these properties are
relevant in and of themselves, i.e., semantically relevant. In the end, Achinstein's proposal boils down to the notion of relevance and degrees thereof. Its originality, if it had one, somehow vanishes.

Achinstein's answer to the second question raised earlier is that "no general answer can be given, but something can be said about various classes of items" designated by A-terms, such as, for example, chemical compounds (1968, p.9). The history of chemistry reveals that before the 17th/18th centuries, the classification of chemical compounds was based on a few physical properties such as color, taste, smell, consistency, solubility, method of preparation, etc. However, in the second half of the 18th century, with the systematic nomenclature of Bergman and Lavoisier, the chemical composition of compounds began to be treated as semantically relevant; and indeed, properties such as color, taste, smell and consistency were generally treated as mere indicators of chemical composition; it was the latter which provided the basis for classification of chemical compounds.

Though Achinstein recognizes quite rightly that the properties associated with a term (e.g., "copper") are empirical discoveries, he holds nevertheless and surprisingly that the association is defensible on linguistic grounds alone (1968, pp.39-42). How could this
be? Concerning the terms here in question (i.e., A-terms), we are in fact never in a position to defend the association on linguistic grounds alone. If by "linguistic association," one means merely the association of a cluster of properties with a linguistic entity, then properties are linguistically associated with the terms here of interest. But if by "linguistic association," one means more, as I suspect Achinstein does, i.e., that the association is a fact of language, a fact by convention, that it is not just an empirical contingency but, given use, an analytical truth that the property is associated with a term, then we come head-on with Quine's strictures against "meaning" and analyticity [1953]. Furthermore, I don't see how this can be reconciled with actual or historical practices. This is surprising in view of the fact that he is well aware of Quine's strictures, and that he seems, unlike his positivist predecessors, genuinely interested in reconciling his account with actual or historical practices. The reason however why Achinstein makes such a claim, is because he (wrongly) believes, like many other philosophers, that it reflects linguistic competence. Thus, for a person to have linguistic competence with a term is, in Achinstein's view, to know and to associate the right cluster of (semantically relevant) properties. 9

In passing, it is pertinent I think to ask why
Achinstein uses the notion of semantical vs nonsemantical relevance rather than simply that of (degrees of) relevance. As we shall see, the reason seems to be that X's semantically relevant properties have presumably something to do with the meaning or rather use of the terms "X" in a way that X's non-semantically relevant properties do not. But as I will also argue this does not establish that linguistic competence requires or presupposes epistemic access to the semantically relevant properties.

So far, we have been dealing with A-terms only, and the general claim that Achinstein has made is that properties cited in a dictionary entry for a term "X" are relevant for X without being logically necessary and sufficient; of these some are semantically relevant, others are nonsemantically relevant. We now need to examine how A-terms are used by scientists. Achinstein proposes the following formulation:

Suppose that as the term "X" is used by a speaker or group of speakers, Pi,...,Pn are properties semantically relevant (both positively and negatively) for X, though not logically necessary or sufficient. Suppose also that they are among the most central ones for X. Then we might have LD2: As the term "X" is normally used by..., items (both actual and hypothetical) are correctly classifiable as X's if and only if they have most or at least many of the properties Pi,...,Pn among others (1968, p.25, my underlining).

This is one way of formulating a (partial) cluster linguistic description for an A-term. It is clearly based on a principle of semantical relevance. In Achinstein's view, LD2 is more realistic, less idealized, and less rigid
than previous descriptions. However, it is plagued with a major difficulty. This difficulty consists in knowing how to apply complex linguistic descriptions of this form to determine whether, given this use, something is correctly classifiable as an X. 10

Achinstein explicitly raises the question whether linguistic descriptions of the use of A-terms can be expressed in a simpler and more precise way (1968, pp. 27-31). He considers various possible models, i.e., in terms of necessary and/or sufficient conditions) and variations of LD2 (introducing numerical considerations, weights and thresholds), which attempt to do just this. 12 But he concludes his analysis quite rightly by stating that, though studying these models can afford us a better insight into the concept of linguistic descriptions, they are either too idealized or too unrealistic or too rigid. They either introduce more structure than actually exists or omit something relevant. But interestingly enough, he does not consider the model that I envisage for the new cluster theory (Part IV).

Now what we need to know is the relationship between a linguistic description such as LD2 for an A-term and the meaning of that term. In other words, we need to know whether when we cite a description such as LD2, i.e., properties which are semantically relevant for being an item
denoted by "X," we are saying something about the meaning of the term "X."

First of all, it is important to sum up what a brief analysis of common usage of A-terms would reveal, namely that the notion of the (a) meaning of a term is not quite appropriate for A-terms, except in special circumstances, for, though in some sense these terms have meaning, they do not have a meaning. Thus, Achinstein, like Wittgenstein (1953), thinks it is better, particularly when dealing with A-terms or rather "cluster terms," to shift our attention to the corresponding question of use. To do so, he argues quite rightly, is not to avoid the issue but instead to put it in a more tractable way. Hence, what we need to know now is whether when we cite a linguistic description such as LD2 we are saying something about the use of the term. It is fair to say that for Achinstein, to acquire knowledge of the semantically relevant properties of the things designated by a term is, in an important sense, to know how to use the term in question. He writes for example:

To know how to use the term "tiger" (in the semantical sense I want to recognize) is to know how to identify tigers in pictures and stories as well as in zoos and jungles. To acquire this sort of knowledge means acquiring knowledge of semantically relevant properties of tigers (1968, p.36; italics added).

Indeed, for Achinstein, to know a cluster linguistic description such as LD2 for a term X involves knowing that as the term is used, properties such as Pi,...,Pn are
semantically relevant for X. In other words, to know how to use a term is to know the cluster of semantically relevant properties associated with it. But what Achinstein seems to overlook is that a semantical phenomenon <semantical use> is "explained" (in a rather circular manner) in terms of a semantical notion <semantical relevance>. Furthermore, he overlooks that someone with such knowledge need not be able to verbalize it, but must be able to use it in classifying items as X's. Or else, s/he must be able to defer to some group of relevant experts or authorities by virtue of the division of linguistic and cognitive labor. This is precisely what Achinstein fails to note and incorporate in his version of cluster theory, and which makes it vulnerable to the criticism later formulated by proponents of a causal theory of reference, and which has often been made against any and all meaning-based or descriptive theories (Kripke, Putnam). It assumes that the individual speaker has some sort of epistemic access, and must therefore know the cluster of properties associated with a term in order to be able to use it successfully. But, as I pointed out earlier and I will emphasize later, a cluster theory need not be committed to such an assumption.

Let us now flip the argument the other way around and assume that someone does not know the semantically relevant properties for X. How would we describe such a person? Can
we say that such a person does not know how to use the term "X"? Given the broad nature of the concept of *use,* (which includes syntactical, pragmatic and semantical aspects), we cannot say so categorically. Clearly, several different situations might be considered which would show that the questions concerning the relationships between semantically relevant properties of \( X \) and the use of the term "\( X \)" are in fact complex ones. Even if someone is ignorant of many or most of these properties, it may be misleading to think that the individual is ignorant of the use of the term "\( X \)," for s/he may know certain other important aspects of its use (e.g., syntactical, pragmatical or even a general category). What we can say is that s/he is ignorant of, or at least does not know very much about, the semantical aspect of use of the term. It seems that this is something that we cannot say about someone who is ignorant only of many of the nonsemantically relevant properties, or of the properties not clearly classifiable as semantically or as nonsemantically relevant. We can therefore infer that to know a cluster linguistic description such as LD2 for a term \( X \) is presumably to know one important *aspect* of *use,* much more so than to know only relevant properties which are not semantically relevant. But, it is important to stress that *not* to know a LD2-type of description does *not* imply that the person does *not* know how to use the term in question.
(see Part IV for details and reasons). 17

To summarize, Achinstein hold that a linguistic description of the use of A-terms will include properties that are semantic- ally relevant and will make rough distinctions among these as to centrality. Achinstein's main contention is that the notion of semantical relevance enables us to determine which properties will or should be cited in formulating a "definition" for an A-term, and also to understand the general nature of the "definitions" proposed in science.

According to Achinstein's analysis, B-terms are defined in terms of logically necessary and/or sufficient conditions. However, the notion of relevance is also useful for B-terms because dictionary entries for these terms may provide information in addition to logically necessary and sufficient conditions. The significant difference between A-terms and B-terms is that LD2 (or a variation thereof) is more appropriate for the former, while a LD in terms of necessary and/or sufficient conditions is more appropriate for the latter. There will be terms on or near the borderline between A-terms and B-terms (e.g., "ion"). On the other hand, changes may occur which might make a term once classified as an A-term classifiable as a B-term now. A term may be treated as an A-term in some contexts, at some point in time, for some purposes and as a B-term in other
contexts, at some other point in time, and for other purposes. The example that Achinstein discusses at some length is the term "acid" (1968, pp. 51-2).

As for C-terms, which are for the most part quantity terms (1968, pp. 54-66), there are typically three types of definitions that scientists give for C-terms: (1) mathematical definitions, e.g., \( k.e. = \frac{1}{2}mv^2 \); (2) operational definitions, e.g., "mass" is defined by reference to operations on a scale as the quantity which two bodies are said to have in equal amounts when they balance each other on a scale; and (3) non-mathematical and non-operational descriptions of the quantity, e.g., "kinetic energy" as "the energy a body possesses because of its motion." Some of these definitions (or combinations thereof) may provide logically necessary and/or sufficient conditions (as in the case of B-terms); others might not, but may provide (semantically/nonsemantically) relevant properties (as in the case of A-terms). Achinstein's analysis in this case concludes that appeal to his notion will not completely determine which conditions will or should be cited in formulating a "definition" for a quantity. For, "not all conditions that are relevant, or even logically necessary and sufficient, are illuminating; some may be quite vague (e.g., "mass" defined as quantity of matter) [1968, p. 66]. His claim here is only that we need
to approach these terms on a case-by-case basis and determine which of the various types of definitions for quantities apply. His concepts might be relevant in deciding which conditions to select for a definition of a quantity and also in understanding the nature of definitions proposed.

To conclude, I think that Achinstein offers indeed a somewhat detailed and fruitful formulation of a version of cluster theory. However, it still faces the brunt of the objections commonly raised against traditional theories of meaning (Part I, end of Chapter 3). First, though "enough" becomes "many" or "most" in his cluster linguistic description, he does not avoid the essentialism (i.e., 3rd grade) which has plagued earlier contextual or cluster theories of meaning—particularly if we take his notion of "semantical relevance" to require irreducibility. Furthermore, though Achinstein sees that the properties associated with a term (e.g., 'copper') are empirical discoveries, he holds that the association is defensible on linguistic grounds alone (1968, pp. 39-42). Second, though he locates the locus of meaning with the (scientific) community, and not with the individual, he assumes epistemic access for the linguistically competent individual. Third, insofar as Achinstein's account assumes—as I think, it does: (i) an inadequate concept-representation, or
conception of meaning (or 'meaning'); (ii) essentialism of meaning (or 'meaning')/an infectious semantic holism, and (iii) a strict determinant relationship holding between a given "(theoretical) background context" and the meaning (or 'meaning') of a term, it will lead, like traditional contextual theories, to paradoxes of extreme and radical incommensurability.

In the final analysis, it is the principle of semantical relevance which is I think most problematic in Achinstein's account. For, as I already pointed out, it reintroduces essentialism --insofar as it requires irreducibility. Furthermore, it is suspiciously circular: for, as we have seen, Achinstein uses a semantical notion to "explain" a semantical question.

In anticipation, I will argue (in Part IV) that it is more appropriate to resort instead to a non-circular, context-sensitive principle, which serves to determine the conditions of assertibility prevailing at some given time for a given community and its members, given some specifiable background knowledge. I shall call such a principle the principle of assertibility. Such a principle could be incorporated into Achinstein's cluster theory in the following manner. Suppose that as the term "X" is used by a scientist or a community of scientists, P₁,...,Pₙ are correlated properties which can be asserted of X; suppose
further that these properties are among the most central ones for X (i.e., have the highest degree of assertibility), then we might have the following partial cluster linguistic description:

LD3: As the term "X" is normally used by a scientist or community of scientists, items (both actual and hypothetical) are correctly classifiable as X's if they have some of the [correlated and assertible] properties P₁, ..., Pₙ, among others. 19

Compared with Gasking's and Putnam's respective versions of cluster theory, Achinstein's theory fares, in my judgement, much better. Though it shares some basic assumptions with both versions of cluster theory, it does not include anything like Gasking's problematic and questionable principle of omnifocality. Instead, it can be made to accommodate something like a principle of correlation (of the properties associated with a cluster-term, whereby some properties turn out to be more central than others). And, as I will argue in Part IV, such a principle can be made to operate at another level as well (i.e., between cluster-terms). Finally, Achinstein's account can also incorporate Putnam's principle of over-/under-determination.
PART II: REFERENCE APPROACHES

In Part I, I have considered among other things various meaning- or descriptive-based accounts, including various traditional versions of cluster theory. I have shown how in some cases they lead to meaning-variance and incommensurability, also how in other cases they simply fail or are inadequate to deal with such problems. I would like now in Part II to examine various reference-based accounts (i.e., those of Scheffler, Putnam, and Kitcher), which are often presented as the only viable alternatives. In anticipation, I will show that they each fail as a viable and adequate account for some reasons or others. In spite of that, one must recognize that they each make significant contributions which must be integrated ultimately by any adequate account.

Chapter 1: Scheffler’s Shift From Meaning to Reference

Scheffler is among those philosophers who have argued that the "paradox of meaning variance" and its consequence, extreme incommensurability, arise only when insufficient and improper attention is paid to the Fregean distinction between sense (or meaning) and reference (1967, p.54-55). In particular, when insufficient attention is paid to the corollary fact that two terms with different meanings can have the same reference and refer to the same thing or entity. He writes:
They are different and it is crucial for philosophical purposes not to confuse them" (1967, p.55). Aside from the contrast in point of clarity, it is important for us to note that sense or meaning (…) differs characteristically from reference. In particular, sameness of sense presupposes, but is not itself presupposed by, sameness of reference (1967, p.56).2

Thus in his *Science and Subjectivity* [1967], he admits that it is plausible enough to suppose, as Kuhn argued, that the meanings of scientific terms might depend somehow on their theoretical contexts, and so might vary with theoretical changes. But he points out, it by no means follows that the references of scientific terms will vary with theoretical changes. A term can have different meanings in different theories and yet still be used to stand for the same thing. Also, he adds:

It is, furthermore, of great importance to realize that we can, and indeed typically do decide issues of reference in specific cases without a prior characterization of sense or meaning in terms understandable to us (1967, p.56).3

Earlier (in Part I), we have seen that the contextual theories failed to come to grips with the paradox of meaning variance and subsequently led to the incommensurability thesis because they lacked an adequate conception of the 'meanings' of scientific terms. They considered scientific terms exclusively from the point of view of meanings, from the point of view of what determines their reference, and directs the use of scientific terms in application to reality. Scheffler's proposal is that we should bypass the obstacles which blocked progress, by simply shifting our
attention from meaning to reference. On the basis of the well-known Quinean strictures against "meanings" in the Two Dogmas (1953, p.47-48), Scheffler attacks the notion of meaning. He writes:

It is ironic in the extreme that this central notion, which is linked to intelligibility, communication, understanding, and significance, is itself wrapped in philosophical obscurity. (1967, p.54). And so he concludes: "Reference is the clearer of the two (1967, p.55).

In support of his recommendation, Scheffler argues that it is concepts of reference, rather than meaning, which are central to our understanding of such basic notions as truth, logical implication, and incompatibility. Suppose we call what a singular term stands for its denotation, and the class of things a predicate stands for its extension. Then we can specify that a singular statement of the form Fa is true just in case the denotation of a is included in the extension of F. Again, the truth of a generalization of the form (x) (Fx ----> Gx) requires just that F's extension be contained in G's. We can go on from this kind of basis to give a recursive specification of what it is for more complex sentences to be true in terms of the referential values of their constituent expressions. Correspondingly, logical relations between different sentences depend on the references of their constituent expressions rather than on those expressions' meanings. Scheffler writes:
As for deductions within scientific systems, it should be especially noted that it requires stability of meaning only in the sense of stability of reference in order to proceed without mishap. That is to say, alterations of meaning in a valid deduction that leave the referential values of constants intact are irrelevant to its truth-preserving character (1967, p.58).

Thus, for instance, one would say that two sentences are logically incompatible if the referential values of their constituents expressions are related in such a way that the two sentences cannot be both true. An obvious and presumably important case is where two sentences respectively attach a predicate and the negation of that predicate to some singular term.

So Scheffler argues that the possibility of meanings inevitably varying with scientific theories is unimportant or irrelevant, or else, not as consequential as it has been presumed.

For identity of reference, as we have noted earlier, does not imply sameness of sense. Terms may denote the very same things though their synonymy relations are catalogued differently. Hence common reference may not only survive alterations of belief outside the realm of synonymies and variations affecting the synonymies of neighboring terms within a given language system. It may also survive synonymy alterations bearing directly on the very terms in questions (1967, p.60).

He adds significantly enough: "[W]hether synonymies are sustained or violated is, in itself, (...) of no real interest for the general purposes of science" (Ibid., p.62). Elsewhere, he urges that "presumed alterations of sense are of no general scientific interest anyway, deductive systematizations in terms of new theories being prized
irrespective of their effects on alleged synonymies of constituent terms" (Ibid., p.64). In the end, Scheffler argues that as long as terms in different theories refer to the same things, then the 'subjectivist argument' from meaning collapses, and we have reinstated "the possibility of disagreement in the sense of explicit contradiction...itself involved in any plausible conception of rational discussion" --since disagreement in this sense requires common reference in order to be differentiated from a mere changing of the subject (Ibid., p.60-1). Furthermore, Scheffler believes that the possibility of constancy of "referential identities throughout variations of theoretical context" lends substance to "the notion of control or justification of theory," and presumably to the objectivity of science (Ibid., p.62).

Let us suppose for a moment that Scheffler is right to insist that it is reference rather than meaning that is of primary importance for the analysis of science. There will still remain a serious hiatus in his approach. For it is clear that an appeal to referential notions will be an effective response to meaning variance and incommensurability only insofar as it can be shown that the referents of scientific terms do not themselves inevitably vary with scientific change, at least with "revolutionary" scientific change. Scheffler states explicitly the
following:

The insistence...that theoretical incorporation affects meanings is plausible at best only with respect to senses, and even so only for certain theoretical incorporations i.e., those considered as altering synonymy-relationships in some way or other. Such alterations, however,...does not automatically effect a disruption of referential constancy nor, therefore, does it automatically disturb the deductive relations which underlie reduction and explanation (1967, p.62).

But Scheffler does not produce any substantial arguments for this coincidence in the referential constancy of scientific terms. It is one thing, to use Frege's famous example, to point to the "Evening" and "Morning" stars to show that terms with different senses or meanings can have the same reference. But it is another to claim that this is the norm, for this is scarcely so. References do, after all, depend (at least in some way or other) on senses or meanings, and it is only in what are in a sense accidentally special cases that terms with different meanings will fail to have different references. It seems quite open for someone of relativist inclinations to maintain that the referential values of scientific terms are as theory-dependent as their meanings, varying whenever the theory in which they are used is modified and eventually changed. The history of science clearly supports such a claim.8

Scheffler does, it is true, suggest that for some terms at least it will be possible for their referential
identities to be determined independently of a characterization of their respective senses by "reinforcement through shared processes of decision on the referential force of a term by application to particular cases." (1967, p.62-4; italics added; this seems, by the way, to be reminiscent of Kuhn's notions of "examplars" and "shared sense of similarity-relations). But it is hard to see how this is supposed to work: Scheffler offers no explicit way to unpack this suggestion. In general, and in accord with the open-ended nature of the process of meaning-making in science, scientific terms will have essentially 'unbounded' referential values: there will always be new and as yet unexamined instances in the future for them to apply to. And this makes it difficult to see how their referential values can be fixed by "agreement on particular cases." We obviously cannot fix (once and for all) the referential value of a scientific term by simply laying down that it applies to each and every one of a given number of identified entities or individuals. If, however, Scheffler's emphasis is supposed to be on the "shared processes of agreement," rather than the "particular cases" agreed on (p.62), then it remains to be shown, in a way that he does not, that such shared processes in general survive theoretical change.9

This brief discussion of Scheffler's proposal does not
constitute a definitive evaluation of the reference approach, but only a beginning toward that end. Such an evaluation would make sense and would be fair only after we have critically examined the proposals, all refinements of the reference approach, made by Kripke [1972; 1980], Putnam [1973; 1975] and Kitcher [1978; 1982]. Thus I turn next to the Kripke/Putnam proposal for a causal theory of reference.
Chapter 2: Putnam’s Causal Theory of Reference

Among philosophers who advocate a reference approach, Kripke and Putnam have made the most stringent and consistent criticisms of contextual theories of meaning, including traditional (descriptive) cluster theories. These theories, as we know, have presented reference as determined or dependent on sense or meaning. The reference (or partial references) of a term is (are) the entities, or things in the world, if any, it stands for. The sense or meaning of a term is a "mentally graspable something" (e.g., a set or cluster of properties or descriptions) associated with the term. And reference is determined or dependent on sense or meaning in the following way: the reference of a term are picked out as that entity, if any, which conforms to the properties or descriptions constituting the term’s meaning.

The theory that Kripke and Putnam propose, known as the "causal theory of reference," represents or rather claims to be a radical but salutary departure from this way of thinking about meaning. In a nutshell, it suggests that the reference(s) of at least some terms are fixed quite independently of any mental senses or meanings (sets or clusters of properties or descriptions) that one may attach to those terms. As Putnam puts it, "'meanings' just ain't in the head" (1973,p.704;1975,p.227).

Leaving aside the case of proper names, let us consider how the causal theory of reference might apply to
scientific terms. Putnam (1975, p.246) extends the causal account to those scientific terms which stand for (i) natural kinds, such as terms used to designate material stuffs (e.g., "gold," "copper," "lead," "water"), (ii) biological species (e.g., "tigers," "elms") and (iii) physical quantities or magnitudes (e.g., "heat," "electric current"). According to such an account, there was an original baptism, involving some sample of material stuff, or some member of a species, or some manifestation (effects) of a physical quantity or magnitude, where the stuff, species, or quantity was dubbed with its name. Here again the story is that any later uses which are causally descended from that original dubbing succeed, precisely by virtue of that descent, in referring to that stuff, species, or quantity.3

Why and how such an account is supposed to solve or dissolve the problems of meaning-variance, incommensurability, and comparison of scientific theories? Kripke himself has not applied the CTR to such problems, but Putnam has (1973, pp.199-209; 1975, pp.245-6). In fact, in his case, these problems seem to have in some sense motivated his development and further application of the CTR.4

As a theory of reference, CTR recognizes that the descriptions held true of the referent r of a term t and
suitable for citation in explaining the meaning or use of a term may vary or change among users of it in circumstances that leave the referential relation invariant. In other words, despite the fact that none of the descriptions applied to a thing or kind of thing apply to it necessarily, and despite the fact that in principle all those descriptions may be rejected, over history, as inapplicable to that thing or kind of thing, nevertheless we may still be talking about, referring to, the same thing. According to Putnam (1973, p.200), one of the major contributions of the causal theory of reference lies in the fact it accounts for the way in which "concepts in different theories may refer to the same thing." The concept of fish, for example, "is continually changing as a result of the impact of scientific discoveries, but that does not mean that it ceases to correspond to the same natural kind" (1973, p.199). Similarly, one could say that the concept of copper (or heat, electricity, electron, field, etc) has changed sometimes radically in different theories, over the course of the history of science, but it still refers to "same things," has the same reference. The relevance of such a theory to the problem of incommensurability and theory-comparison must be obvious.5

As a preliminary however to Putnam's "solution" to the problem of incommensurability/theory comparison, one must
also note some of the assumptions underlying such a solution. (i) It agrees with the extreme "Kuhnian" thesis that scientific theories, across a revolutionary change, are incommensurable and therefore incomparable in terms of the "meanings" of the terms involved, or else in terms of their descriptive vocabulary. Whether these "meanings" are characterized in terms of clusters of properties/descriptions rather than sets, does not presumably improve the situation. (ii) The continuity between such theories must be found elsewhere than in the "meanings" or descriptive vocabularies. (iii) The only alternative left to look for and explain such continuity is in the reference of a term. We must therefore reject any account which presupposes some essentialism of meaning, and adopt instead an account which is based on an essentialism of reference.6

Thus, in Putnam's causal theory, what a scientific term refers to need have nothing to do with what the users of the term currently think, believe or know (possibly erroneously), nothing to do in other words with what theoretical assumptions involving that term they presently adopt (perhaps wrongly). The references of their terms are not picked out as those things or kinds of things which satisfy those assumptions, but simply as those things or kinds of things which were involved in the baptismal or
inaugural occasions when the terms were first introduced. So the problem of deciding which part or sub-part of the theory fixes reference disappears, since no part or sub-part does. Also, the reasons for doubting that scientific terms have determinate referential values recede and somehow vanish. Indeed, in such an account, there seems to be no remaining reason for doubting that scientific terms can have the same references: all that is required is that the use of those terms stems from the same causal origin for the proponents of both theories.

The causal theory of reference faces however a number of difficulties and objections. Once they are examined and answered, it will become apparent that it cannot fulfill the promise of any real or easy "solutions" to the problems here of interest. I would like to raise systematically some of the difficulties and objections faced by the CTR, and then go on to examine them in more detail.

(1) Putnam's way of solving or dissolving the problems of incommensurability/theory-comparison is (a) to accept their existence as far as anything we can say about a thing is concerned, and (b) to shift from the idea of "meanings" (or descriptions) to that of "reference" which is alleged to be independent of meanings or descriptions. My first question concerns the alleged meaning- or descriptions-independent reference. Does Putnam, and Kripke for that matter, argue
for this claim in a convincing way? I have some serious doubts which I will vent in a moment.

(2) Can a CTR explain or even make plausible the continuity of scientific discussion? Can it be of any help in explaining how scientists can be talking about the same thing in different theories? How are we to decide that the reference is the same? Presumably, we must trace the use of the term in question to some "baptismal or inaugural event" causally connected to the present use. But why should such a causal connection through time be taken as confirming common reference? Does (and can) a CTR offer a satisfactory explanation in this respect? My initial assessment leads me to believe that it does not and cannot -- at least as it is formulated by Putnam. [But about later improved formulations of CTR by Kitcher]. In anticipation, it is worth pointing out that the claim by CTR that successive and potentially radically different descriptions have a common reference which establishes continuity seems to be just an arbitrary assumption. Is not such assurance of continuity or commonality of reference an empty promise in most cases? What about those cases where we have a theory in which the central concept or term (and the corresponding entity or kind of entity) is abandoned and replaced by a new theory, with a new concept or term (and a new sort of entity corresponding to it. I am thinking of such cases as
phlogiston versus oxygen theories, ether theories versus special relativity, caloric versus kinetic theories of heat, etc --not to mention all other cases involving less radical and drastic changes not only in meaning but in reference as well. This brings out still further the inadequacy of a Putnam-like account. It also seems to suggest that there is far more involved in the commensurability and comparability of scientific theories than the mere continuity of the sort focused on exclusively by reference theories (causal or otherwise).8

Let us now return to point (1) (and subsequently (2) mentioned above) and examine in greater detail the claims of the CTR. Let us suppose with CTR that an initial or inaugural baptism need not tie necessarily and/or sufficiently any (number of) specific descriptive criteria to the term that is being introduced. It still remains that this does not tell us what type is being named. It is one thing to say that the reference of "copper" (or "gold" or "lead") is fixed as whatever was originally dubbed with the name, quite independently of what anybody might believe that referent to be like. But this claim is quite unspecific without some further information of what type of thing was there being named. What needs to be specified, it seems to me, is at least that it was a material stuff that was being named, and not, for example, a tool, a color, a coin, or a
statue. Even though my arguments have been confined thus far to stuff terms, they can easily be extended and applied to species and physical quantity terms.9

The question that I have just raised is that of the nature of our concept of material stuff. What is it in other words for different samples to be samples of the same stuff? According to Putnam and Kripke, the requisite similarity is being similarly composed out of the basic "constituents of matter" (Newton-Smith, 1981, p.173). Thus, they might suggest that what makes all samples of copper (or gold) examples of the same stuff is simply that they are composed similarly with respect to their constituents atoms, electrons, etc.10

Doesn't this kind of reasoning put an end to the relevance of CTR for the problems here of interest? For what happens when different groups of scientists have different ideas about the constitution of some material stuff? If what stuff terms stand for depend on such ideas, then surely we will once more be without any guarantee that the same things will answer to different assumptions involving a stuff term in different theories. Indeed what assurance will we have that there will be anything in reality answering determinately to any of their various clusters of assumptions? To take a well-known example, in the early part of this century there were disagreements as
to whether the term "lead" applied only to the stable matter with atomic weight 206, 207, or 208; or whether it applied as well to the radioactive substances with atomic weights 210, 211, 212, or 214. The protagonists here seem to have had different ideas about what is required for something to be of the same stuff as the samples originally dubbed "lead" (which presumably were predominantly composed of lead 206, 207, and 208) [see Asimov, 1985; Leicester, 1952; Partington, 1961]. But then surely CTR implies that if they were using the term "lead" to refer to anything, they were using it to refer to different categories of matter.

A proponent of CTR could with some justice complain that this line of reasoning misses in fact the point. It is not what scientists take to make something the same stuff as the original sample that fixes the reference of a term like "lead." To suppose this is to fall back into the traditional way of thinking which had reference depend on the sense or "meanings" users attach to a term. It is rather what it is to be the same material stuff as the original sample that fixes the reference of a term like "lead." That different groups of scientists might at different times have different ideas about the properties of lead is neither here nor there, from CTR's point of view. All that this shows is that those groups (except perhaps one) are wrong about what lead is like, and consequently,
they are in danger of mis-identifying it. Recall that an account of how scientific terms refer is designed to deal with questions about semantic-ontological objectivity, and not primarily with epistemological objectivity questions. How scientists or groups of scientists are best to choose assumptions involving the term "lead" for example, and how they are best to identify samples which satisfy those assumptions, are epistemological questions. In a CTR, it seems that we are concerned instead with prior questions of what those assumptions are about, or what does satisfy "lead." CTR postulates that something must be preserved for referential stability and continuity, i.e., something "essential"; but it denies that the required preservation is epistemic. We do not in other words stabilize a term's reference by associating certain properties or descriptions with continued used of the term.11 It is rather the world, independently of our linguistic and theoretical predilections, which dictates the conditions for retention of reference. The essence on which reference depend is metaphysical; it need not be an epistemically or experientially accessible feature of the referent.

This defense of CTR would be a valid response to the objection as originally formulated. However, the objection here question can be reformulated and made to bite at a deeper level. For scientists do not differ on just such
specific questions as what physical constitution is characteristic of lead, as opposed to other stuffs (e.g., copper, gold). They can also differ in their general concept of material stuff itself, i.e., on the general question of what it is for two bits of stuff to have the same physical constitution. 12 In ancient times, the Greeks took the basic physical entities to be the four elements of fire, earth, air, and water. Dalton, in the 19th century, took them to be the indivisible atoms of the various chemical elements (Partington, 1961). During the 20th century, we have come to recognize an expanding range of sub-atomic particles, and their number keeps growing. Differences on this level naturally give rise to differences in notions of what it is for two bits of matter to be the same stuff. And such differences do undermine, it seems to me, the causal account of natural kind terms.

Suppose for a minute, as CTR holds, that specific ideas about what differentiates lead for example from other substances play no part in fixing the reference of "lead." It seems to me that the general notion of what it is to be a given substance will nevertheless still play a crucial role. In fact, it is only on the basis of some such general notion that one way rather another of "generalizing" from the original sample is laid down. And so when different groups of scientists have different such notions, they will
extrapolate from such original samples in different ways: for each group the larger entity that was exemplified in the original sample and named in the baptism will come out differently. It is quite possible then for theoretical differences about the principles which identify kinds of material stuff to result in different groups of scientists using a term like "lead" (or "copper," or "gold") to refer to different things. Couldn't (shouldn't) we regard the controversy about "lead" in the early part of this century in just this light?

The development of scientific theories about the sub-atomic structure of the atom in this century raised the question of what exactly was sub-atomically required for two bits of matter to be the same substance. In particular, there was the question of whether it required atoms with exactly identical nuclei or whether similarity merely with respect to the nuclei's positive charges would suffice. The latter answer would indicate, if my understanding is correct, that the appropriate generalization from the "original sample" of lead for example was to the category comprising all bits of stuff with the same chemical properties as that sample. The former answer, on the other hand, would require in addition that the bits of stuff in the category extrapolated to also shared atomic weights with the original sample (cf: Leicester, 1952). Thus, we would
have here an example of a general difference on stuff identity-principles leading to a difference in the reference of a specific stuff term.

At this point of my argumentation, I think it will no longer help for a proponent of the CTR to complain that it does not matter what scientists take to make two things the "same stuff," that the important point is what it is for them to be the "same stuff." Now, unlike previously, we are left with no further story to tell about what such questions are supposed to be about in the first place. The main point of the CTR, in my view, is that it can explain how the reference of "lead" for example could be fixed independently of any (theoretical) assumptions involving the term in question, independently of any meanings or descriptions associated with it. Presumably, the original sample, at the initial or inaugural baptism, plus the concept of "same stuff" serve that purpose. But, as I pointed out, nothing is said about what fixes the references of "same stuff" --if it is not assumptions (meanings, descriptions) involving that term. Consider again the issue behind the "lead" controversy: is it right to think of classes of things with chemically identical atoms as the "same stuffs," or only classes of things with mechanically identical atoms? The problem here is not just how to decide this question. It is rather and most importantly for my present purposes, that we
have been given no account of what it is about, no real or satisfactory reason to think that those who gave different answers were really talking about the same thing.13

Most likely, in the final and ideal (scientific) account of reality, if there is ever one, there will be various principles for using the terms "same stuff," "same substance," or "same basic physical constitution." But can we without further argument assume (like the CTR does) that our present use of such terms is such as to make them determinately answerable to the facts they will stand for in the final and ideal theory?14 According to the CTR, we are supposed to have been referring all along to what lead (or copper, or gold, or electrons) really are, as opposed to what we think or may have thought (possibly erroneously) they are. Can we say for example that what the Greeks would have meant by the "same substance" would have made coal, graphite and diamonds all the "same substance"? Insofar as their ideas about composition in terms of the four elements had any significance, why should these not be counted as "different substances"? After all, their different crystalline structure does mean that coal burns, graphite rubs off, diamond shines, etc. Furthermore, is it obvious that the 19th century notion of "same substance" was such as to make the isotopes of chlorine, or lead, the "same substance"? If anything, 19th century chemists required
identity of atoms for similarity of substance. Now, if as these examples show, we cannot presuppose that past uses of "same substance" answered to the same things as our present use of this term, then why should suppose that either past or present uses answer to the same things as the final one will?

The possibility of variation and even change in assumptions about stuff identity principles means that the CTR cannot assure us that scientific terms refer to the same things as used in different theories. For the sake of argument, it may nevertheless suggest a way of getting some distance from the holism of "contextual theories of meaning." We can for example construe stuff identity-principles as at least playing a privileged and central role in the fixing of references of scientific terms. But this would not be acceptable for the original proponents of CTR (i.e., Kripke and Putnam) --even though it might be for later proponents of the reference approach (e.g., Kitcher, 1978; Newton-Smith, 1981). Such a construal would clearly concede that reference depends at least in part on what (theoretical) assumptions are accepted. And the CTR would thus be undermined in its major claim. However, if we did allow that kind identity assumptions played a special role in fixing references, then the original contextualist and holistic argument for doubting that scientific terms can have determinate references would
lose force and fall by the wayside, and there would be the possibility of common references whenever such assumptions are shared. However, not even this suggestion stands up to a close-up examination. In order to see why, we need to consider further what is involved in the adoption of identity principles for some type kind, material stuffs, species, or physical quantities. [In the forthcoming discussion, I shall be concerned generally with the different types of natural kinds].

To begin with, let us assume that a given natural kind consists of sub-entities (bits of lead, copper, tigers, amounts of heat), each of which share a cluster of properties. In theorizing about a natural kind we will correspondingly have a number of generalizations stating the various properties shared by most (perhaps all) of the examples of that kind. These properties might include both (i) manifest, common-sense, observable, pre-theoretical properties, such as color, shape, size, tactile impression, etc, and (ii) underlying structural properties, like atomic constitution (material stuffs), or genetic make-up (species), or basic physical activities involved (in quantities).

One direction in which we expect to be able to push our theories is towards conceptions of "underlying structure" which will allow the various manifest properties
of each of the different kinds of some type to be explainable by reference to their characteristic underlying structures. And Putnam does just that in several places (1975, pp. 235, 239). He writes:

[W]e know that there are kinds of things with common hidden structure but we don't yet have the knowledge to describe all those hidden structures (1975, p.244). If there is a hidden structure, then generally it determines what it is to be a member of the natural kind, not only in the actual world, but in all possible worlds. Put another way, it determines what we can and cannot counterfactually suppose about the natural kind (1975, p.241).

Thus we would expect the various observable properties of copper (or lead) and its observable interactions with other substances to be all explainable in terms of its basic physical constitution. In a sense, the notion of "characteristic underlying structure" is a notion of the sort of property which will account for all the manifest properties of the kind in question. Insofar as our scientific theories do seem to capture such underlying structures, then of course we will take the various kinds of some type to be differentiated (in essence [?]) from each other by such characteristic underlying structures.

What this clearly indicates is that as we develop our ideas, our scientific theories, there will be an interaction between assumptions about what manifest properties show something to be an example of a given kind and assumptions about what underlying structure is characteristic of that
kind. On the one hand, we will seek an idea of underlying structure which will account for the recognized cluster(s) of manifest properties.15 On the other hand, we will want to tailor our views on which manifest properties indicate membership of the kind to our existing and current conception of its underlying structure. This tendency leads to the rejection of such erroneous assumptions as that "all that glitters is gold" (fool’s gold) and "all carbon is black" (what about diamonds?). But the former kind of aspiration is of course what leads to our developing new ideas about underlying structures in the first place [e.g., as when we try to identify the molecular constitution of compounds, or when we opt for atomic number (or nuclear charge) rather than atomic weight as showing what counts as "lead," "copper," or "gold"]. Thus, in general terms, it seems that it is the former kind of aspiration which leads to the adoption of certain "kind identity-principles" rather than others. We choose just those ways of differentiating kinds in terms of underlying structure which promise to yield systematic explanations of the currently known and admitted cluster(s) of manifest properties.

In the end however, there are two ways in which we can change our criteria for identifying examples of given natural kinds: (i) we can alter our ideas of which manifest properties indicate membership of the kind, or (ii) we can
alter our ideas of its underlying structure. Which kind of revision is to be made at any juncture will depend (at least in part) on the relative methodological worth of the alternatives. Is empirical progress more likely to follow from adjustments in manifest criteria, or from a different account of underlying structure? In general, the first option will seem more appropriate, for the theoretical revisions involved in changing manifest criteria will be far less central than those required by altered conceptions of underlying structures. Let us recall in passing that part of what makes them count as "underlying structures" is that many generalizations about manifest properties can be explained by reference to them. However, there will also arise occasions when more fundamental revisions are called for, when new ideas about underlying structure will in the end lead to a better account of the acknowledged clustering, or correlation among manifest properties. Thus, it would be a mistake to characterize the scientific process (and its goal) as one in which ideas of underlying structure should always predominate and take precedence over ideas about which manifest properties indicate natural kind membership. Indeed if it were never appropriate to stick with existing ideas about manifest properties, there would never be any pressure to change our ideas about the underlying structures that differentiate natural kinds.
Yet it is precisely this kind of proposition, that assumptions about underlying structure should always play a special and privileged role and predominate over manifest properties, which is presently under scrutiny. If assumptions about underlying structures did play a special and privileged role in fixing reference, then the referents of natural kind terms would always be guaranteed to be such as to make those assumptions true, and it would always be a mistake to reject them. This proposition does, I must concede, get some apparent support from the fact that it is usually appropriate to discount previously accepted ideas about manifest properties when they conflict with assumptions about underlying structure. But, as I pointed out, it makes no sense to suppose that it is always appropriate or even right to do so. If it were, then nothing would ever lead to our developing new ideas about underlying structure. If the history of sciences is any indication, scientists have as a matter of fact developed radically new ideas of underlying structures, and there is no good reason to believe that they would not continue to do so in the future, if and when judged appropriate (cf: Shapere, 1984, chap.18).

My line of argumentation thus far leads me to conclude also that it would be wrong to give any special and privileged role to "original samples" in fixing the reference of natural kind terms. The point of introducing
"original samples" at the inaugural stage (or "dubbing ceremony") was in a way to get reference fixed independently of any accepted assumptions, apart may be from abstract assumptions about identity principles for the type of kind being named. In particular the idea of "original sample" in the CTR was supposed to preclude assumptions about the manifest properties of a kind from playing any part in fixing the reference of the term for that kind. But we can now see that in doing this the introduction of "original sample" destroys at the same time any possibility of accounting for changes in our ideas about underlying structures and kind identity-principles. If what a kind term referred to was just whatever shared "underlying structure" with a given "original sample," and if assumptions about how to recognize examples of the kind were never of any import in fixing this reference, then it would follow I presume that whenever our view of underlying structures conflicted with such assumptions about manifest criteria or properties, it would be the former that was true and the latter assumptions that were false. If, for example, our ideas about reddishness/yellowness, malleability, ductility, conductivity, combinatory properties, etc were really of no importance for what counts as a bit of "copper," beside the requirement that it shares "underlying structure," according to some conception thereof,
with a particular sample of stuff, then we ought always to discount as examples of "copper" any substances, however apparently similar, which are not, according to our current conception of "underlying structure," of the same structure as that "original sample." But, as we have seen earlier, we do sometimes change our conceptions of "underlying structure," precisely because doing so enables us to better account for the manifest properties of "copper" (and other material stuffs).

In the end, the idea of "original samples" for natural kind terms has somehow always been somewhat implausible. For, if it implies that the original sample is all that counts, and associated clusters of properties are irrelevant for reference fixing, it is difficult to see how such a term could ever change its reference. This ought to be in principle at least possible.16 As Kitcher, for example, is quick to note in his attempt to improve upon the CTR, the references of scientific terms do as a matter of fact change. In response to this kind of objection, some defenders of the reference approach have loosened the idea of "original sample," allowing in effect that the relevant sample can be expanded by later actual applications of the term, and perhaps contracted, for example when it is decided that some earlier applications did not fit the term's usage after all. The main question to be confronted here is: what
governs these expansions and contractions? How are they made? Presumably one would have to say that they result from accepted assumptions (theoretical and otherwise) about how to recognize and classify instances of the kind in question. If that is so, there seems to be no remaining reason for not rejecting the causal reference approach altogether, and reverting to a (new and revised) contextual/cluster-based theory of meaning. According to such a theory, insofar as a given natural kind term refers to anything at a given time, it refers to just those things which have the cluster of properties, manifest and underlying, that current theory ascribes to the kind.

Before closing this discussion, it is I think useful to bring in Putnam's hypothesis of the division of linguistic labor (1973, p.209; 1975, pp.227-8;245-6). [In light of Kitcher's recent article (1990), it might as well be referred to as the hypothesis of the division of cognitive and linguistic labor]. According to Putnam, in our use of many terms we defer to "experts," or authorities on a given subject within a particular domain of knowledge and inquiry. We allow that what such terms refer to depends not on the ideas (assumptions, descriptions or meanings), such as they are, which we as individuals associate with those terms, but on how the cognoscenti (i.e., those who know) use those terms. Consider for example the term "gold"
(or "copper" for that matter). As Putnam points out (1975, p.227-8), very few people really know what in the end shows something to satisfy this term. Yet all of our utterances and judgements involving this term are answerable to what a "special sub-class of speakers" would count as satisfying it. A similar point can be made presumably with respect to proper names.17 Putnam characterizes the universality of the hypothesis as follows:

Every linguistic community exemplifies the sort of division of linguistic labor just described; that is, possesses at least some terms whose associated 'criteria' are known only to a subset of the speakers who acquire the terms, and whose use by the other speakers depends upon a structured cooperation between them and the speakers in the relevant subsets (1975, p.228; see also pp.227-8;245-6 for further elaborations)

However illuminating this hypothesis of the division of linguistic and cognitive labor is, one should not consider it, as is often the case, as an argument in favor of the CTR, or else as intrinsically bound to a CTR account. For, it is perfectly consistent with any other account of what makes a term as used by experts refer to what it does. Since it does not show in particular that the descriptive facts that expert speakers associate with a given term are totally irrelevant to what that term refers to --even if the associations of inexpert speakers may be so irrelevant, I will therefore re-consider in a more favorable light the case of cluster theory as a viable and adequate account, which is based among other things on what I shall call a
"principle of deference" (to experts or authorities).

I hope to have shown that since there is no remaining reason for not rejecting the CTR altogether, there are good reasons to revert to a contextual/cluster-based theory. While I am contemplating taking up this task in Part IV, it seems only fair to the reference approach to consider the proposal made by Kitcher [1978; 1982] which presumably offers refinements and improvements of the CTR.
Chapter 3: Kitcher’s Context-Sensitive Reference Approach

Though a proponent of the reference approach, Kitcher is critical of other proponents of such an approach (e.g., Scheffler, Field, and Kripke/Putnam). Furthermore, he takes seriously the claims made among others by Kuhn and Feyerabend about the relevance of the history of science to philosophy of science. As a result, as we shall see, he proposes a theory of reference based on the notion of "reference potential" which is context-sensitive and which allows for reference-changes. But, as I will show, it faces its own problems and difficulties.

Kitcher begins his article "Theories, Theorists, and Theoretical Change" (1978) by noting that:

Paradoxically enough, the writers who have contended most vigorously that history of science is relevant to philosophy of science have also argued for theses which imply that the task of the historian of science cannot be successfully completed (1978, p.519).

He is referring to Kuhn, (1962,1970) and to Feyerabend (1962, 1964). It is correct that they have exhorted philosophers to take the history of science seriously, but have they also argued for theses, which necessarily imply, according to Kitcher, that the task of the historian cannot be successfully completed? ‘That sometimes the content of past theories resist expression in modern terms’ makes the task of historian difficult and challenging, but not impossible. According to Kitcher,

Historians of science are interested in discovering what
[a given scientist, say] Priestley was talking about, and how much of what he said is true. Their researches are relevant to philosophy. For it is a commonplace of modern philosophy that our views about the nature and development of science can be illuminated by studying the processes through which the great dead theories were replaced by their modern counterparts; and, since our only access to the great theories is through the writings of the great dead theorists, we can only engage in such study if we can decide what those theorists were talking about (1978, p. 519, my addition in brackets; italics added).

It seems quite reasonable to say that historians of science are interested in finding out what a given past scientist was talking about. However, as Kuhn has repeatedly pointed out, including very recently (1988 Lectures), it is not as uncontroversial to say that they are also interested in finding out how much of what s/he said was true.

Now, is Kitcher talking about intra-theoretical truth, or rather inter-theoretical, and theory-independent truth? Or simply about truth from our vantage point? It seems that for Kitcher, one need not make such distinctions, nor give any reasons for such a claim. There can be only one truth, of course, which successive scientific theories approximate. This may be the best assumption to make in order to explain and make sense of the scientific enterprise, but it remains to be established that this is so. Also, it remains to be shown that an alternative account without such an assumption cannot equally make sense of the scientific enterprise without any significant loss. In the meantime, let us note
that the cluster theory that I envisage ultimately will constitute such an alternative (see Part IV). Given the debate and controversy about realism and anti-realism in science, this should be an advantage rather than a deficiency or failure.1

In line with Kitcher's interpretation however, if Kuhn and Feyerabend are right, then the historian's enterprise is impossible; and we cannot formulate past theories in contemporary language. In order to disprove this, Kitcher proposes a "sophisticated strategy" combined with a sensitive reading of the history of science for understanding the semantical aspects of theoretical change. His goal is to show how such a strategy solves or dissolves the problems raised by Kuhn and Feyerabend, and thus makes the task of the historian of science possible. To further make his case, he applies it to a particular example, one of Kuhn's favorites: the revolutionary development in the 18th century which led from phlogiston theory to oxygen theory.

Before turning to this strategy and ultimately evaluating its merits, it is only appropriate to examine beforehand Kitcher's formulation of the problem as he sees it. In Kitcher's view, both Kuhn and Feyerabend have argued, despite disagreements on points of details, that during the major upheavals in the history of science, "scientific revolutions," scientists of different
Theoretical persuasions do not have recourse to a common body of observational evidence or to a shared set of methodological rules. More fundamentally, they are unable to communicate. Both sides lack the ability to express, within their own language, the assertions of the rival theory. The *preconditions* for *rational* debate between them breaks down. Kitcher refers to this view as the doctrine of *conceptual relativism*. More precisely, it is "the doctrine that the language used in a field of science changes so radically during a revolution in that field that the old language and the new language are not *intertranslatable*"—in a truth-preserving way, Kuhn would add (1978, p.520).2 Despite the careful attention paid to well-known examples in the history of science (e.g., Ptolemaic/Copernican astronomy, Priestley's phlogiston theory/Lavoisier's theory of combustion, Newtonian dynamics/Einstein's special theory of relativity), the main idea of radical difference in meaning across the revolutionary divide has remained unclear. Thus, according to Kitcher, "conceptual relativism inherits the philosophical difficulties of the notion of meaning" (or sense, to use Frege's expression), which have been exposed repeatedly over the years (cf: Quine, 1953).

Like many other philosophers who have felt uncomfortable with the notion of meaning or sense, he urges
us to recognize the benefits of doing as much semantics as possible with the notion of reference. He writes in this respect:

If we are to clarify the doctrine of conceptual relativism, it is natural to turn to the notion of reference, reformulating the doctrine as the thesis that, for any two languages used in the same scientific field at times separated by a revolution, there are some expressions in each language whose referents are not specifiable in the other language (1978, p.521; italics added). And he adds further on:

What would block understanding --and what is of interest to Kuhn and Feyerabend-- is a situation involving a special type of referential change, namely change which culminates in a mutual inability to specify the referents of terms used in presenting the rival position. Cases of this kind do appear to threaten the possibility of an objective comparison between the rival theories and hence to subvert traditional accounts of intertheoretic debate (1978, p.522; italics in text).

Instead of arguing for referential stability or continuity for that matter (contrast with Scheffler and Putnam), and instead of arguing that the theses leading to incommensurability and incomparability are absurd or self-defeating (like Achinstein, Davidson, Kordig), Kitcher argues rightly that we should attend to the possibility of formulating the opposing theories in the rival languages --even though reference may change. It is at this point that we need to proceed with a more sophisticated approach to the reference of scientific terms than is usual in philosophy of science, otherwise the historical evidence will indeed support conceptual relativism.3

The theory of reference that Kitcher has in mind is a
"context-sensitive theory" (CST), rather than a "context-insensitive theory" (CIT), which is what philosophers of science usually envisage when constructing a theory of reference for a particular language. Briefly put, the latter theory aims to correlate expression-types of L1 (e.g., Aristotelian language) with expression-types of L2 (i.e., contemporary English), and to complete matrices of the form "In L1, e refers to. . . ," where e is an expression-type, and where the blanks are filled with an expression of L2 which is co-referential with e]. The assumption here is clearly that we can find a mapping of expression-types (of L1) onto expression-types (of L2) and that all tokens of the same expression-type (of L1) can be treated in the same way. But, as Kitcher is quick to point out, such CIT's are inadequate for dealing with natural languages and the languages of the natural and social sciences, in which different tokens of some expression-types refer to different entities in virtue of their production in different contexts. Instead, he argues, these languages require a CST. But what is a CST? And in passing, how different or similar is it to Putnam's CTR? And what will it enable us to do?

According to Kitcher, theories of reference for particular languages must, as one expects, meet standards of adequacy which are laid down by what he calls the general
theory of reference (1978, p.524). Such a theory provides us presumably with universal principles for the determination of reference, which we accept independently of our views about the referents of expressions in particular languages and to which we appeal to evaluate such views. Though Kitcher does not offer a detailed and explicit version of this theory, he claims that CST's will be adequate theories of reference for the languages mentioned above: they will specify the referent of tokens of some expressions-types of a given language (the context-sensitive ones) by invoking general principles about reference. In effect, he is assuming quite rightly that we have some tacit or intuitive views about the principles constraining determination of reference which it is the task of the general theory of reference to elaborate and codify.

One such principles which constrains translation and reference-determination, and which historians typically use, according to Kitcher (1978, p.534), is the principle aptly dubbed by Richard Brandy "the principle of humanity" (1973, p.492-452). [In passing, let me note that the cluster theory that I will envisage in Part IV will also incorporate such a principle]. This principle requires us to impute to the speaker whom we are trying to understand or translate a "pattern of relations among beliefs, desires and the world [which is] as similar as possible to ours" (Brandy, 1973,
p.443). It is a constraint because it compels us to recognize that the intentions of the speaker on the occasion of the utterance play an important role, and they ought to be identified only in relation to, and together with what we know about the speaker's environment and behavior, and our current best understanding of the pattern of relations among beliefs and desires, on the one hand, and the world, on the other, which is shared by human beings. Kitcher claims to have adopted at least tacitly such a principle in his ascriptions of reference in his case-study. But he warns quite rightly that if we attempt to satisfy this principle, and still treat all tokens of the same type in the same way, then we shall be led to the position defended by Kuhn and Feyerabend," i.e., incommensurability and incomparability of scientific theories. Presumably, we can avoid such a consequence, while still satisfying the "principle of humanity," only if we recognize as he does that the tokens of the same type need not be treated in the same way.5

Kitcher outlines some of the other key ideas and principles of his general theory of reference. Like other philosophers who advocate the reference approach (e.g., Putnam, Kripke), Kitcher supposes that the general theory of reference is a "historical explanation" theory. And that as such, it enables us to "gain a clearer view of the issue of conceptual relativism and of the semantical aspects of
theoretical change" (1978, p.527). However, as we shall see later on, it is unlike the traditional version of the causal theory of reference in a number of respects. For example, and interestingly enough from the point of view of my limited goal in this project, Kitcher seems to suggest that it may be compatible with a descriptive (or meaning) approach, at least to some extent. On the other hand, it claims to account for the fact that a scientific term may have a number of referents and that these referents may change over time with theoretical change. Kitcher writes:

The general theory of reference that I espoused above proposes that the referent of a token of an expression is the entity which figures in an appropriate way in a historical explanation of the production of that token (1978, p.537; italics added).

Clearly, this characterization is imprecise. We do not know what is meant nor what is to be understood by "figures in an appropriate way in a historical explanation." Kitcher gives us however a rough idea as follows:

An explanation of the production of a token will consist in a description of a sequence of events whose final member is the production of the token and whose first member is either an event in which the referent of the token is causally involved, or an event which involves the singling out by descriptions of the referent of the token. Let us call the first member of the sequence the initiating event for the production of the token (1978, p.537; my underlining).

Elsewhere, in a 1982 paper "Genes," in which he reformulates his theory and re-applies it to another particular case, Kitcher states more explicitly:
An event is the *initiating event* for a token if the hypothesis that the speaker referred to the entity singled out in that event provides the best explanation for her saying what she did. And explanations are judged by their ability to provide a picture of the speaker’s intentions which fits with her environment and history and with the general constraint of the Principle of Humanity (1982, p.347; italics added).

This notion of *initiating event* bears some similarity to Putnam’s idea of an *introducing event* (1973, p. 203), but it differs from the latter in that Kitcher is explicitly concerned with cases in which various events are or may be associated with one expression-type. Further assumptions in Kitcher’s view include: (i) Different tokens of the same type can be linked to the world in different ways. (ii) The linkages between scientific terms and the world are constantly renewed. Being in accord with the open-ended nature of the process of meaning-making in science, these assumptions are unobjectionable and must be reckoned with by any adequate theory.

Suppose we want to provide a theory of reference for the language used in presenting some past scientific theory. Kitcher claims that there are in general four possible outcomes:

1. We can find a CIT adequate for the language under study.
2. We cannot find a CIT which is adequate for the language under study. We can find an adequate CST, and, using the CST and the available evidence, we can specify the referent of each token produced by a speaker of the language.
3. We cannot find an adequate CIT. We can only find a CST and, in the light of the CST and the available evidence, there are some tokens produced by speakers of the language whose referents we cannot specify. However,
for each expression-type of the language we can specify a set of entities such that the referent of any token of that type belongs to the set. 

(4) We can only find a CST, and, for some expression-types we are unable even to specify a set of entities such that the referent of any token of that type belongs to the set (1978, p.528).

According to Kitcher, philosophers of science have typically assumed that the only way to respond to the kind of conceptual relativism that presumably Kuhn and Feyerabend brought upon us is to argue that all scientific language can be treated as in (1). They have thus assumed that scientific terms are context-insensitive. In contrast, he argues instead that the apparently problematic parts of the language used by past theorists can be treated as cases of type (2), and at worst, such languages provide cases of type (3). But such cases as described in (3), otherwise referred to as cases of underdetermination of translation, they do not preclude the possibility of comparing two scientific theories; rather, they require that each scientist considers a number of different ways of formulating his rival’s position, but each of these may be compared with his or her own position. As for cases of type (4), which articulate the idea of radical incommensurability and incomparability of two scientific theories, and which presuppose that there is no adequate translation of the language used in presenting one of those theories in the language of the other, Kitcher claims they do not occur that often, relatively speaking.
If, as Kitcher suggested above, conceptual relativism is intended to be the first part of an attack on the thesis that scientific theories can objectively be compared, and if, the central claim of conceptual relativism is the thesis that, in general, scientists working in the same field but separated by a revolution, find, when they attend to one another’s languages, that they are confronted by cases of type (4), the relativist must first establish this central claim in order to contend plausibly that incommensurability and incomparability obtain. In other words, only if the conceptual relativist establishes this central claim will s/he be able to argue that, lacking any language in which to formulate their disagreements, scientists separated by a revolution are unable to compare their respective theories. But, if any of the other cases obtains, Kitcher argues, the scientists in question will be able to formulate their disagreements, even though they may not always be able to resolve them. But for Kitcher, this is somehow beside the point, since he is only “concerned with the entering wedge of the attack on the objectivity of scientific decision, i.e., the thesis that the preconditions for debate are not satisfied” (1978, p.528). I would tend to agree with Kitcher in this respect: these preconditions are satisfied in most cases for the most part, and the problem of incommensurability or incomparability of scientific theories
is less threatening than has been presumed.

Given his characterization of the general theory of reference as outlined above, Kitcher claims that we can explain how different tokens of a scientific term (or the same expression-type) can refer to different entities. Also, he claims, his approach accords better with the view that the linkage between scientific terms and the world is constantly renewed. Concerning the first point, it suffices to suppose as he does that the production of different tokens can be initiated by different events. As for the second point, we have to consider the further development of his account of reference for scientific terms. In this respect, he writes:

When we construct an explanation of the production of a token, we attempt to link that token to an entity in the world through an initiating event, and (...) our construction aims to make comprehensible the judgements and inferences of our subject. To say that an entity 'figures in an appropriate way' in the explanation of the production of a token is, I suggest, to claim that the hypothesis [that the token was initiated by an event in which that entity was causally involved or singled by description] best explains why our subject makes the assertions and arguments he does. We might say that the initiating event is that event whose effects are manifested in the subject's linguistic behavior when he produces the token (1978, p.538; italics added).

Alternatively, we might say that we identify the initiating event as the event such that the hypothesis that our subject is referring to the entity involved in that event best explains why s/he says the thing s/he does. In Kitcher's terms:
The task is to use our understanding of the pattern of relations among mental states and the world which is common to humanity, together with the available data about the speaker's (or writer's) environment and behavior, to identify the intentions which were operative on the occasion of utterance and thus construct an explanation of the production of the tokens produced (1982, p.343).

And this is presumably what a seasoned historian (or philosopher functioning as such) does or should do, even if s/he does not or cannot always formulate explicitly what s/he is doing. Again, as Kitcher puts it:

The best explanation of the utterance tells the correct story about the speaker's intentions in making the utterance, relating those intentions to the external circumstances of the utterance and to the speaker's verbal behavior (1982, p.343)

In some cases however, Kitcher points out, our search for the initiating event may lead us back through events involving other speakers to some "act of baptism" by the first user of the term (cf: Krikpe and Putnam's CTR). However, when we are concerned with the utterances of scientists who are developing particular theories and using the vocabulary peculiar to those theories, he recognizes that different explanations will often be appropriate. We may find for example that a scientist's argument presupposes that the referent of a term satisfies a particular description or cluster of descriptions, and we may best explain his or her utterances by hypothesizing that they are initiated by an event in which the referent of the term is fixed by that
description or cluster of descriptions.\textsuperscript{7} If "historical explanation" accounts of reference (like CTR) often assume that there is a single chain connecting tokens of a term to the object singled out on the first occasion of its use, Kitcher is suggesting a different and rival picture. In the latter, the connections of terms to the world are often extended in subsequent uses. This picture, one must concede, reflects better the continued reapplication and redefinition which is typical of scientific usage.

Now the question however is whether this kind of account of the ways in which scientific terms refer threatens itself the possibility of a common scientific language, of comparability between scientific theories, and of communication among scientists. In Kitcher's view, it does not. In order to quell our concerns in this respect however we need to attend to further aspects of his account. He suggests that

\textup{[A]}n expression-type used by a scientific community is associated with a set of events such that production of tokens of that type by members of the community are normally initiated by an event in the associated set. The set which is associated with a particular expression-type (in a particular community) will be called the \textit{reference potential} of the expression (for that community) \textup{[1978, p.540; italics in text].}

Elsewhere, he characterizes the \textit{reference potential} of a term-type for a given linguistic or scientific community as "a compendium of the ways in which the referents of tokens of the term are fixed for members of the community" \textup{[1982,}
This notion of referential potential is perhaps the most significant contribution of Kitcher (see also Kitcher, 1982, pp.345;347; and compare and contrast with my "field of referential possibilities" in Part IV). It is akin to Frege's sense insofar as the latter is defined as "the manner in which the reference is presented." By identifying scientific concepts with referential potentials, we can clarify, Kitcher claims, the idea that theoretical concepts must absorb theoretical hypotheses and so enhance our understanding of conceptual change in science (1978, p.543). Thus, we may call terms which have a heterogeneous reference potential, that is, with two or more different initiating events, theory-laden terms. Their use depends upon hypotheses (explicitly or implicitly formulated by the community of scientists) to the effect that the same entity is involved in the appropriate way in different events which belong to same referential potential. The nature of the dependence is such that if one of the hypotheses were to be seriously questioned then the use of the term which depends on it would have to be revised. Can we avoid using theory-laden terms? Kitcher believes quite rightly that we cannot. In fact, he argues that the framework that he is providing can enable us to elaborate the famous Hempelian dictum: "concept formation and theory formation go hand in hand" (Hempel, 1966). Does then the use of such terms pose
problems for scientific communication, and ultimately for theory-comparison? Kitcher also believes that on the whole it does not. In fact he believes that "the thesis that scientific term-types have heterogeneous reference potentials is the key to solving the problem of incommensurability" (1982, p.346). He explains as follows:

Because the reference potential of scientific expressions is frequently heterogeneous and presupposes the generalizations of a particular theory, it will often be the case that there are some crucial expressions in the languages of theorists, separated by a large change in theory, whose reference potential cannot be matched by any expression of the rival theory. Nevertheless, successful communication can continue, even when reference potential has changed, because each theorist can specify the referents of his rival's individual tokens (1978, p.547; italics added).

It is important however to note that Kitcher does not pretend that we can always decide which event initiated the production of a given token, or that we can always identify the reference potential of a given expression-type. But he does suggest that we can often determine changes in the reference potential of a scientific term, and that the claims of historians about conceptual change are best understood as changes in reference potential. As he puts it,

The notion of reference potential is a tool for exposing the fine-grain of the history of science (1978, p.546; italics added).8

To summarize: Kitcher's account is a "historical explanation" account of reference similar in a number of respects to Kripke's and Putnam's CTR, but different and more plausible. Unlike the latter, it allows for the
reference of a scientific term to change over time, and it somehow incorporates some aspects of the descriptivist (or meaning) approach. Recall that the referent of a token in Kitcher’s account is the entity which figures in an appropriate way in the (correct) historical explanation of the production of that token. The production of the expression-token is the terminal event in a sequence of events which would be described in detail by the correct and complete explanation of that terminal event. And this sequence links the expression-token produced to an entity singled out in the first causally involved event of the sequence, and that entity is the referent of the token.

But, as Kitcher is quick to note, in some (quite a few) cases, the referent may be singled out by a description or cluster of descriptions in the initiating event; in such cases, he suggests that the referent be represented as determined by the description or cluster. As a result, we may have to use a cluster of associated descriptions, instead of a set of events, in order to characterize and exhibit the ways in which the referents of tokens of an expression-type are determined. Berent Enc has argued convincingly that some scientific terms (i.e., theoretical terms such as phlogiston, caloric heat, kinetic energy, id, entropy, etc) have their referents fixed by description (1976, pp.261-282). Insofar as the referents of different
tokens of some scientific terms are fixed via events in which different yet correlated descriptions are used, we might then have to represent the reference potential of such terms as clusters of descriptions. Kitcher's account needs, it seems to me, this extension or provision in order to be able to deal with theoretical terms (e.g., electron, phlogiston, etc). Now, would this suggest that Kitcher is indeed making some serious concessions to the descriptivist or meaning approach (CTR's main competitor and arguably an alternative to it)? Despite what he says, I think that he is indeed making such concessions. He claims that, if he is making such concessions, they are in fact only apparent and superficial. This is because, like many other philosophers, he has in mind the traditional descriptivist approach. For example, he claims that the description or cluster of descriptions which determines the referent of a speaker's token is not necessarily one which the speaker is supposed to know and provide; rather, it is only a description or cluster of descriptions used to single out an entity in the event which initiated the production of the token. Does this qualification suffice to make his account different from a descriptivist account? Kitcher's assumption here is that such an account requires necessarily that the speaker be able to know and provide the appropriate description which serves to pick out the referent. But as I shall argue
in Part IV, this assumption is not necessary for a descriptivist account, particularly for one which we may call in anticipation a "social-historical descriptivist account." For such an account, like the CTR (or variation thereof), can well accommodate the idea that the speaker need not know what the appropriate description or cluster thereof is, which serves to pick out the referent of a term; the speaker may even only have a false description, and yet manage to refer to whatever entity is referred to by the appropriate and correct description or cluster thereof. In so doing, the speaker in question would be however deferring to an expert or group of experts on the matter (cf: Putnam's and Kitcher's hypothesis of the division of linguistic and cognitive labor, and my suggested corollary the principle of deference).

Getting back to the main thread of this discussion, Kitcher attempts to show further that his approach can account for how the development of science can change the reference potential of a term or expression. In a given domain of research, scientists may make discoveries and come to accept new hypotheses. This may lead to the absorption within the reference potential of a term of a new class of events or cluster of descriptions through which the reference of tokens of the term can be fixed. Thus, if for example some scientists come to believe that circumstances
of different types confront them with the same entity, or if they become convinced that the entity which satisfies one description (or cluster thereof) satisfies another, then a term whose previous usage was initiated by one type of event or by one of the descriptions may undergo an expansion of its reference potential. Alternatively, as in the cases of scientific revolutions, scientists may learn that what they had believed to be causal interactions with the same entity are in fact encounters with different entities, or that descriptions taken to be coextensive are not coextensive. Previously accepted hypotheses can be disconfirmed. These kinds of circumstances may lead instead to a contraction of reference potentials. Rather significantly, he writes:

When we put both types of change together, we see how radical conceptual change is possible as the result of a continuous process. At a later time, the reference potential of a term may share no common element with the reference potential at an earlier time. Yet, the evolution of the concept may be continuous, in that, at intermediate times, the reference potential of the term may contain elements from both the non-overlapping classes (1982, p.346).

Elsewhere he claimed that by bringing these two kinds of changes together, his approach will be able to account for "radical conceptual revision without conceptual discontinuity" (1978, p.544). And he goes to develop the following scenario for how this can happen: the reference potential of a term may be first expanded through the addition of a new cluster of events or descriptions, and
later on scientific progress may lead to a contraction of the reference potential through deletion or abandonment of all except the newly added cluster. In fact, there need not be deletion or abandonment, but simply reconfiguration and restructuring of the relevant and appropriate cluster, whereby those events or descriptions which were previously held to be central and determinant are now secondary and derivative while a new cluster of events or descriptions is considered central and determinant. [It should be obvious from the above that a new and revised cluster theory will be able to incorporate these insights]. Kitcher concludes his 1978 proposal thus:

[O]ur approach promises to solve a problem which has bedevilled most accounts of conceptual change in science, the problem of expressing the idea that while scientific concepts can change radically, they also change continuously (1978, p.544).

The diagnosis is unquestionably correct, but can Kitcher cash in on his promise? In the end, I think that Kitcher’s proposal faces some serious problems, even though once again it constitutes a significant improvement on the Kripke/Putnam CTR. I will consider two such related problems in the abstract, and go on to examine another set of problems which arise particularly when we apply Kitcher’s proposal to an example in the history of chemistry --one of Kuhn’s favorites and one which Shapere considers to be a test-case for any adequate theory.
Problem 1: In the end, Kitcher seems to be suggesting that we will be able to provide an account of the phenomena of conceptual change by "charting the shifts in referential relations between words and the world" (1982, p.339); in other words, by charting the changes in reference potentials of scientific terms. His main assumption, as we have seen, is that scientific terms frequently have reference potentials which include diverse initiating events. But one needs to ask: what binds or connects these various initiating (or reference-fixing) events together? Suppose a satisfactory answer is to say that it is the continuity (or evolution) of conceptual development in science, whereby tokens of a term-type are applied and re-applied in an attempt to arrive at a more and more adequate characterization of certain entities of interest in a given field of scientific research. Then, it seems to me, we need to stress that these initiating or reference-fixing events must be actual and not just fictitious; fictitious initiating events will not do. Even though Kitcher has taken great pains to outline clearly his approach, it is difficult in practice to specify and to trace the various initiating events that scientists periodically use to connect a term-token to a given phenomenon. I presume that ultimately in Kitcher's account term-tokens must be organized into trees branching out, and even form semantic or lexical networks (cf: Kuhn, 1988). Every once in a
while, a particular term-token is connected to a particular event. The tokens may change their structure, the character of the initiating events might even change, but in the face of such change, the sequences of application and re-application can be traced presumably through time. Assuming that their continuity is guaranteed by their causal relation which makes the system work, there will still be a problem persisting, namely, how do we partition conceptual trees of term-tokens into types? To my knowledge, we are not given a method or even hint for a criterion. Clearly, isolated term-tokens are not sufficient for communication or description. But even connections of successive term-tokens will not do. Understanding language in general, and scientific language in particular, requires both term-tokens and term-types and some way of relating them. In anticipation, this will also represent a problem for the new cluster theory that I envisage, but only one with a solution --at least partially (Part IV).

**Problem 2:** As we have seen, link-to-link sequences of term-tokens do change their referents through time. Ways of fixing reference also do change. Also, we have seen that Kitcher assigns a role to individual language users and the communities they form. His characterization of such role is certainly useful and even plausible. However, it faces a serious problem because of the way in which Kitcher
delimits or delineates his linguistic community. Kitcher writes:

With respect to a particular expression-type, two speakers belong to the same linguistic community if they are disposed to count exactly the same events as initiating events for production of tokens of the same type (1982, p.346). [A] linguistic community, with respect to a term, is a set of individuals disposed to admit the same initiating events for token of that term" (1982, p.347; italics added).

According to Kitcher, the only agreement necessary for scientists to belong to the same community is agreement over initiating events. This agreement can presumably come about in many different ways. On the one hand, the speakers might both be aware of all the events they would count as initiating events for the term-type, so that their agreement could be established in principle by having them describe the relevant events. On the other hand, they might have no clear conception of how their references are fixed, and simply rely on some third party, so that they would concur in virtue of a general disposition to defer to the references of an expert in the relevant domain. There might also be various kinds of cases between these two extremes. Thus, Kitcher avoids presumably all the variations which might be due to differences in beliefs about the "meanings" of the term at issue. In this connection, Kitcher emphasizes two points:

First, two members of a linguistic community may differ in their beliefs. What is crucial is that they agree on the ways in which the referent of a term should be
fixed. Second, not all community shared beliefs which use a particular term may be employed in fixing the reference of that term. It is quite possible that each member of a linguistic community should be prepared to assert that the things referred to by a particular term lack a particular property, and yet use that term to refer to entities which have that property. So long as the belief is not used to fix reference, a false belief may prevail throughout a community (1982, p.347).

Hence, in Kitcher’s account, members of the same linguistic community can disagree about all sorts of things, just as long as they agree about the way in which the reference of the term at issue is fixed as well as the initiating event that fixes its reference. Each term-type delimits its own linguistic community, generating thus as many linguistic communities as there are term-types. Despite this however, these communities might be coextensive for large clusters of terms.

The main problem however is that Kitcher’s linguistic communities do not coincide with scientific research groups or communities as we know them in actuality. If scientific research communities can be described in terms of cooperation and competition, then scientists can belong to the same community even though they are not in total agreement on a variety of issues, including initiating events for the production of term-tokens of the same type. For the most part, scientists belonging to the same research community will be in substantial agreement over a variety of issues, including the initiating events for many of the term-types they use, but an occasional
disagreement on this score does not automatically exclude a scientist from his or her research community, as Kitcher seems to suggest. For Kitcher, linguistic or scientific communities are individuated solely for the purposes of fixing reference. As such, they are to some extent social fictions. It would have been more convincing and plausible if Kitcher were able to define scientific research communities in terms of the relevant and actual social relations which characterize them. In passing, it should be noted that a scientist can belong to several research communities of increasing inclusiveness, but restrictions on time and opportunity limit the number of communities to which any one scientist can belong at the same time.

I would like now to consider briefly Kitcher's proposal as it is applied to the 18th century chemistry, i.e., the episode in the history of this science which has led to the "chemical revolution," the demise of phlogiston theory and its replacement ultimately by Lavoisier's oxygen theory.

According to Kitcher, "it may be useful compare the descriptions of these familiar reactions given by the phlogiston theory and by modern elementary chemistry" (1978, p.530). And so he sets up the following table:

<table>
<thead>
<tr>
<th>Phlogiston Theory</th>
<th>Modern Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In his account, Kitcher argues rather convincingly that the language of 20th century modern chemistry can be used to identify the referents of the terms and expressions of 18th century chemistry (i.e., phlogiston theory), at least to the extent that those terms and expressions actually refer. By taking us through the texts of Priestley and Cavendish, while inviting us to think of their experiments in modern terms, Kitcher claims some obvious identifications can be made: "dephlogisticated air" sometimes refers to oxygen itself, sometimes to an oxygen enriched atmosphere. As for "phlogisticated air," one can see, he claims, that it is air from which oxygen has been removed. Also, we might come to see that the expression "x is richer in phlogiston than y" is coreferential with "x has greater affinity for oxygen than y." Furthermore, in some contexts, the term "phlogiston" in the expression "phlogiston is emitted during
combustion," does not refer at all, while in other contexts, it refers to hydrogen (see Kitcher, 1978, pp.530-536;545).

Depending on the kind of history that they are interested in, historians (and philosophers) of science looking into the texts of past scientists can and must use modern language and something like Kitcher's context-sensitive strategy in their attempts to identify the referents of past scientific terms. In this sense, a context-sensitive reference-determining strategy may provide historians and philosophers of science with a stock of unproblematic concrete examples from which they might then attempt to learn what the problematical terms and expressions in past scientific texts mean or refer to. But does this strategy yield, as Kitcher claims, a translation? What kind of translation? A truth-preserving translation? Does such a strategy bring all discussions of incommensurability to a close --as he seems to suggest?

By involving the introduction and use of modern terminology, Kitcher claims that his strategy may serve to explain why and in what areas or respects past scientific theories were correct or incorrect, true or false. Thus, he would say for example that statements about the substance emitted upon combustion were false, while statements about the effect of dephlogisticated air on vital activities (respiration) were true because in those statements
"dephlogistrated air" referred to oxygen. It seems to me that Kitcher is only using his strategy in conjunction with modern theory to explain why and in what respects past scientific theories were successful or failures, i.e., why some statements made by proponents of a past scientific theory were confirmed, others not. Such a task is certainly part of any attempt at interpreting past scientific theories and the texts in which they are formulated. But, as Kuhn correctly pointed out, interpretation or even Kitcher's strategy does not allow us to say that some individual statements containing terms from a past scientific theory are true or false. Scientific theories, it seems, must be considered and evaluated as holistic structures --at least regionally or locally. There is certainly intra-theoretical truth, but it is not clear, in light of the historical evidence, how we can make sense of inter-theoretical truth --which Kitcher assumes without argument (Kuhn, 1982, p.17).

Returning now to the questions raised earlier, let us contemplate for a moment what a past scientific text/theory "translated" by Kitcher's strategy would look like. It is not clear for example how non-referring occurrences of "phlogiston" would be rendered. As Kuhn pointed out, one option would be to leave blank spaces. This seems to be suggested by Kitcher's silence on the subject and by his concern to preserve truth-values, which, as we have seen, is
problematic. But what does it mean to leave blank spaces in a "translation"? Given the standard philosophical use of this term, one can but say that it is not yet a translation or else that it is a failed attempt at translation. If only referring terms or expressions possess translations, then many past scientific theories could not be translated because they might contain non-referring terms or expressions. The texts containing past scientific theories report what past scientists believed, independent of its truth-value --in the sense assumed by Kitcher, and it seems that a translation must communicate exactly that (Kuhn, 1988: Part III).

Presumably, Kitcher would use the same strategy for referring terms like "dephlogisticated air." On this contrual, "phlogiston" would then sometimes be rendered as "substance emitted from burning bodies," sometimes as "metallizing principle." and sometimes by still other expressions. But as Kuhn argued, this strategy is problematic not only with terms like "phlogiston" but with referring expressions as well. Assuming that the use of a term like "phlogiston," together with a compound term like "dephlogisticated air" derived from it, is one of the ways by which the original text communicated the beliefs of its author, what would obtain if we substitute unrelated or differently related terms or expressions for those related,
sometimes identical terms or expressions of the original text? We would most likely end up suppressing those beliefs and producing an incoherent text. That is why Kuhn stated that if we examine a Kitcher "translation," "one would repeatedly be at a loss to understand why those sentences are juxtaposed in a single text" [as schematically presented above] (1982, p.7).14

In general, I think that Kuhn's point in this respect is well taken. If many, perhaps most, terms in older scientific languages and theories are identical in form and function with terms in a modern language or theory, others are new and must be learned or relearned together, as a components of a Cluster of correlated terms. These are the untranslatable terms for which the historian or philosopher has to discover or invent meanings in order to render intelligible the texts in which they appear. But interpretation (or language-learning, as Kuhn later suggests [1988]) and not translation, is the process by which the meaning or use of those terms is discovered or invented. Translation does not seem to arise. Thus, I agree with Kuhn when he says that the task or activity of the historian or philosopher seeking to understand past scientific texts is perhaps better seen as falling under "hermeneutics," i.e., the science (and art) of interpretation" rather than as "translation" as Kitcher and many other commentators have presumed (see Part IV: interpretative principles of the new
cluster theory).

All in all however, I think that Kitcher has produced a rather insightful account. But as we have seen, it faces some serious problems and difficulties, and as a result, it is far from being satisfactory.
PART III: RECENT DEVELOPMENTS IN PHILOSOPHY OF SCIENCE

In this part, I will review and evaluate the views of Shapere (1984, 1989), Nersessian (1984, 1987, 1989) and Kuhn (1982-90). Despite differences and disagreements on specific issues, these views share a lot in common --a lot more than is acknowledged by these authors themselves. For example, they all reject the reference approach as a viable avenue for an adequate theory of the process of meaning-making in science. They also reject the cluster approach --even though some commentators have attributed such an approach to each of them. They have all adopted a thoroughly historical and cognitive approach. They make room for the "context of discovery" in their respective account of the process of meaning-making in science; they concentrate on actual cases of scientific change and development.1

And interestingly enough, they all argue --including Kuhn, that the problem of incommensurability is in fact less threatening to the objectivity and rationality of science than has been assumed; they claim that there are different kinds and degrees of incommensurability, and that, for the most part, it is only "local incommensurability" within a larger background of commensurability which occurs. Thus, on these views, incommensurability does not lead necessarily to untranslatability, incommunicability or incomparability.

This being said however, it does not mean that there is
no problem for philosophers and historians of science to contend with. The point is that we do translate, interpret and even understand different scientific theories; in short, we do manage to compare them despite differences between them of various kinds and degrees (conceptual and otherwise). The question then is: How do we do that? Which theory can better provide an account of how we can and do compare different scientific theories? Which theory can better accommodate by the same token the historical evidence often adduced as a sign of incommensurability and as a cause for alarm. I think it is about this question that differences and disagreements between Shapere, Nersessian, and Kuhn are most pronounced—even though, once again, they overlap in many respects. For example, they each formulate what they take to be the minimal requirements for an adequate theory of the process of meaning-making in science.

As we shall see in forthcoming chapters (of this Part III), they argue for different approaches: (1) Shapere argues for a thoroughly historical approach, one which I shall call "a reason-based transtheoretical approach"—presumably neither meaning nor reference based.2 (2) Nersessian urges "a cognitive-historical approach," i.e., one which stresses the need not only for an appropriate concept-representation but also for greater attention and concern to the processes by which concepts are altered,
modified and subsequently abandoned. As we shall see, her approach owes much to Shapere’s, at least with regards to concept-representation, to her emphasis on reasons, reason-related criteria for "defining" a concept, her emphasis on the continuity and evolution in science despite discontinuities by way of "chains-of-reasoning connections" etc. However, her distinctive contribution lies in the cognitive dimension of her approach; for, more so than Shapere, or Kuhn for that matter, she insists that there is much to be gained by philosophers of science interested in the problem of conceptual change if they reach out for insights (theoretical and methodological) from the cognitive sciences, in particular cognitive psychology and artificial intelligence.3 (3) As for Kuhn, he seems in this recent and unpublished work of his, to adopt "a locally holistic verification theory of meaning."4 His approach is clearly historical --at least in one sense of history, as we shall see-- as well as cognitive; but it is based on semantical notions (drawn from linguistics)5 such as lexical structure, hierarchies of taxonomies, feature spaces, semantic fields and contrast sets, etc. It is interesting to note that 28 years after the publication of his 1962 work, Kuhn is still attempting to overcome some of the difficulties, ambiguities, terminological confusions, and exaggerated implications which have plagued his work ever since.
Before turning to a more detailed examination of these views, I would like to point out that, in some sense, each of the views here in question can be considered as exemplifying a revised version of cluster theory, as I envisage it, despite explicit denials by Shapere [1989] and Kuhn [1988]. In effect, they are "closet cluster theorists." They deny that they are cluster theorists because, I think, they only have in mind a particular version of cluster theory, and that is, a traditional version. But, as I shall argue in Part IV, with appropriate revisions and adjustments, a cluster theory can accommodate and integrate the insights and contributions that they have made respectively; in particular, I will argue that it can satisfy the requirements that they each formulated for an adequate theory of the process of meaning-making in science. Finally, I will point out systematically in what respects the new cluster theory differs from the accounts offered by the authors here in question.
Chapter 1: Shapere’s Reason-based Transtheoretical Approach

In order to situate Shapere’s approach, it might be helpful to examine its supposed connection to Putnam’s. In his 1973 paper “Explanation and Reference” Putnam stated that

with a few possible exceptions (e.g., Feyerabend), realists have held that there are successive scientific theories about the same things: about heat, about electricity, about electrons, and so forth; and this involves treating such terms as ‘electricity’ as transtheoretical, as Dudley Shapere has called them (in 1969; reprinted in 1984), i.e., as terms that have the same reference in different theories...The main technical contribution of this paper will be a sketch of a theory of meaning which supports (...) Shapere’s insights. (1973, p.200; italics in text).

Indeed, it is in this paper that Putnam laid down much of the groundwork for his later work on the problem of incommensurability and theory-comparison. Even though he later abandons some of the ideas presented therein, he would retain the idea of developing a theory on the basis of the notion of "trans-theoretical terms." Even though Putnam claims to sketch a theory of "meaning," he in fact develops what would come to be known as a causal theory of reference, as we have seen in Part II, chap.2. His main contention is that a theory of reference as opposed to a theory of meaning is the key to solving the problem of incommensurability and that of theory-comparison. Even though Putnam claims to have been following Shapere’s insight, that of a "transtheoretical term," Shapere disagrees with the
interpretation that Putnam gives to his notion.

In a recent paper, Shapere writes:

He [Putnam] understood me as having meant by this expression "terms that have the same reference in different theories." That is not at all what I had in mind, at least not in the sense in which Putnam interpreted the role of reference --as being the key to understanding how scientific theories might be said to be talking about the same things. On the contrary, the idea I was proposing when I wrote of "theory-transcendent concepts" was, and remains today, very far from the interpretation placed on it by Putnam (1989, p.425-6).

In order to bring out the distinctiveness of Shapere's view, I will draw out some of the similarities and differences between the main approaches discussed in previous Parts. In Part I, we have seen that according to contextual theories of meaning (e.g., Carnap/Kuhn), the meanings of scientific terms are determined by the theoretical context in which they occur. This view led, as I have tried to show, to the problem of how such theoretical contexts can be compared. In Part II, we have seen that Putnam's strategy in the face of such a problem was to resort to the concept of reference as a means of dealing with it. But as I have tried to show the reference approach cannot help us in dealing with that problem. According to Shapere, there is another alternative, which he has in mind in discussing "theory-transcendent" or "transtheoretical" terms. Is it truly an alternative, a reasonable and viable one? We shall only raise the question at this point, and attempt to answer
it later on, after we have had a chance to present more fully Shapere's view.2

Shapere's alternative view accepts certain features of both the "meaning" and the "reference" approaches, while also rejecting certain aspects of each. In accord with the meaning approach, it holds that all we can say about electrons is given by the (cluster of [?]) properties (descriptive terms) ascribed to them; also, in agreement with the more recent contextual theories within this tradition, it holds that the reasons for ascribing those properties (descriptive terms) to electrons are provided by a larger "background" context. On the other hand, however, in opposition to the meaning approach and in agreement with Putnam's reference approach and his criticisms or objections against the above mentioned approach, Shapere maintains that none of those (clusters of) properties or descriptions are essential in the sense that they must be present in order for anything to be an electron. Shapere agrees with Putnam, for quite different reasons, as we shall see, that any of those properties or description can be rejected, modified, or replaced in light of what we learn as the process of scientific inquiry takes place. The alternative view advocated by Shapere focuses on properties or descriptions rather than reference; it is thus a meaning approach (perhaps of the cluster kind), but it does not consider any
of those properties or descriptions to be distinguished as part of the "meaning" of the term in question. But if the uses of terms differ when those uses occur in different theoretical contexts, does not Shapere's view face at this point the old problem of theory-comparison? How can they be compared?3

According to Shapere, it is in fact precisely at this point that his view differs most fundamentally from both the "meaning" and "reference" approaches. "For on my view, he writes, continuity of reference is established by there being reasons for changing the body of properties ascribed to an entity or type of entity" (1989, p.427; italics in text).4 He makes this point more precisely as follows:

If, for two usages (whether of the same term or of distinct ones) U1 and U2, respectively determined by theoretical contexts T1 and T2, there is a chain-of-reasoning in terms of which we can understand why certain properties ascribed in usage U1 and its successors up to and including U2 were abandoned, altered, or replaced, then that chain-of-reasoning connection explains the possibility of comparing the two usages and their theoretical contexts despite the fact that, within each, the usages of the term or terms involve few or even no properties in common (1989, p.427; my underlining).

Though Shapere mentions the usages over time of the same term or of distinct terms, he explicitly considers only cases involving the former. In contrast, I will consider both kinds of cases (see Part IV, Tables 2 and 3). Interestingly, Shapere's main point in the passage quoted can be illustrated by means of a model, which reveals the
affinities of his view with that of Wittgenstein (1953) and his notion of "family resemblances."

<table>
<thead>
<tr>
<th>Usage Ascribed</th>
<th>Theoretical Properties</th>
<th>Context</th>
<th>Essential (none)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U1 T1</td>
<td></td>
<td>ABCD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U2 T2</td>
<td></td>
<td>ABC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U3 T3</td>
<td></td>
<td>ABCF</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U4 T4</td>
<td></td>
<td>ABGF</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U5 T5</td>
<td></td>
<td>HKGF</td>
<td></td>
</tr>
</tbody>
</table>

To be more accurate however, and in accord with the open-ended nature of the process of meaning-making in science, one should perhaps add further perpendicular connections (---) after U5, T5, and HKGF. The similarities to Wittgenstein's "family resemblances" must be obvious. But Wittgenstein said very little about what the resemblance-relations were like, except that they were resemblances. In contrast, Shapere sees particular usages
as being related to one another by reasons (i.e., by chain-of-reasoning connections), and thus as important for understanding the development of our world-view.

One may however object to Shapere's model by pointing out that its simplicity fails to exhibit some of the complexities of scientific change. In addition to the ambiguities surrounding the notion of "theoretical context" of a particular usage, a variety of "reasons" are adduced for scientific changes. For example, the reasons for changes in usage of a particular term may lie in the alteration or rejection of other parts of the theory than those having to do directly with that term itself; there are a variety of types of such reasons. Indeed, Shapere recognizes that there are other grounds or reasons, besides direct comparison of properties ascribed in two different contexts, for judging a new scientific development to be continuous or discontinuous with something that has gone on before; and some of those other bases can affect judgements of whether "the same entity" is being spoken of in the new context. For example, it is very likely that changes in a list of conservation principles (of matter or energy) can affect such issues. But since Shapere's concern is with the problem of comparison of concepts in different theoretical contexts, he focuses only on continuities and changes having to do directly with property ascriptions.
Thus, despite the simplicity of his model and the possible objections raised, Shapere claims that the basic idea of his model still holds—at least in some cases, and that a more elaborate treatment could deal with the possible complications (see new models of cluster theory, Part IV, Tables 2 and 3). He then turns to an application of his model in order to show how it might work, in contrast with both the traditional "meaning" and "reference" approaches. For this purpose, he chooses the term "electron" and bases his analysis on the historical account offered by A. Pais, *Inward Bound*, (1986).

The term "electron" has passed through a series of historical stages. In Stoner's view, who has rethought Faraday's electrolysis experiments, an electron was an indivisible unit of charge (not of matter). Thomson in turn who has found charge to be invariably associated with a definite mass, decided to add that property to the property of charge, and to consider an electron to be a classical particle—a mass, with position and velocity, that is, momentum, and possessing negative charge. However, with the advent of quantum mechanics (with Bohr, for example) and the multitude and variety of reasons advanced for it, the simultaneous attribution of position and momentum were abandoned, and other properties, such as spin, added. The quantum mechanical concept of the electron was carried further in quantum field theory, in which electrons,
considered to be a variety of leptons, are like all particles, i.e., excitations of the quantum field possessing a range of properties (quantum numbers) that Stoney or Thomson or even Bohr could not have conceived. According to Shapere, all these theories were "about" electrons -- not because some essential properties were ascribed to electrons, not because some set of necessary and/or sufficient conditions of application was given or even because some cluster of conditions of application of the term was given, nor, as Putnam would have it, because we were presumably "referring all along to the same thing." In fact, Shapere finds that the basis of such an assertion (an unclear principle of benefit of the doubt)9 has been left unjustified, and is in fact unjustifiable. Note also that he explicitly rejects a cluster approach -- even though he does not explain fully his reasons.10 He writes instead that:

[The unity of the "electron tradition" is provided by the chain of reasoning that led to the successive additions, deletions, and modifications in what was said of electrons, from Stoney to the latest developments in quantum field theory. (1989, p.428 italics added)].11

But Shapere adds:

If we look only at the extreme ends of the historical process -- at what Stoney held about electrons, and at what contemporary particle physicists say about them -- we will find those usages, if not radically incommensurable, at least impressively different (1989, p.428; my underlining).

As we shall see in the next chapter, Nersessian (1989)
makes a similar point on a different example ('field'), as if to suggest that the traditional approaches have failed because they have only adopted such a perspective. For Shapere, as for Nersessian, it is nevertheless easy (in principle, if not in practice) to see how and why (in the sense of reasons, not causes) we got from there to here in our thinking about "electrons" or "fields." Under the scheme that they have in mind, the differences between extreme episodes in the development of a concept, and the comparability of these extremes would be explained. Shapere writes explicitly:

In short, the continuity of the inquiry and the theories resulting from it -- the legitimation of the assertions that they were all inquiries about the electron and that they were all theories of the electron -- is provided by the chain-of-reasoning connection linking the successive phases of the history. The term "electron" has a transtheoretical use, above and beyond the theoretical uses that occur at specific stages of the history, when it is applied to that entire reason-linked series, as it can be despite the in-principle possibility that the theories constituting the beginning and later stages of the history of electron-research may ascribe non-overlapping sets (or rather clusters) of properties to electrons (1989, p.429; italics in text; my addition in parentheses).12

Thus, if I understand Shapere's view correctly, continuity and comparability are not given either by common ascriptions of properties or by common reference. Instead, the rationale for saying that we have referred to the same thing all along is given by the chain-of-reasoning connections (CDRC's) which gives continuity to the history.
What distinguishes Shapere's approach is that on his account both changes of properties or descriptions ("meanings") and commonality or continuity of reference are explained, as he puts it, "by looking at science not in terms of its linguistic aspects (which is presumably what Putnam's reference approach does), but as a process of inquiry in which reasons and evidence are provided for whatever changes are introduced" (1989, p.429). For, he explains,

"It is here, in the reasoning employed in science, and not in the philosophy of language, that we ought to have been looking all along if we want to understand the nature of science and its development (1989, p.429; italics and underlining added).

Shapere has used the "electron" example to illustrate the continuity of scientific discussion, the issue with which Putnam's causal theory of reference was designed to deal, and with which it failed to deal successfully (see Shapere, 1982, pp.1-23 for a detailed discussion of his objections and criticisms of the CTR). His discussion was also intended to show that the issue in question can be dealt with successfully via quite a different line, i.e., his chain of reasoning approach. The question then is: can Shapere's approach deal as successfully with other important sorts of scientific change, which are not captured by the electron example? Shapere claims that it can.14

In order to show that, he proposes to single out a particular sort of scientific change. The sort of
scientific change he has in mind is not only interesting and important in-and-of-itself, but it brings out still further the inadequacy of a Putnam-like account. It involves "cases in which science abandons a particular term central to a theory, and thus abandons the entity to which the term refers" (1989, p.429). Shapere states:

Perhaps the most important thing we want a theory of scientific change to be able to do is to show how an older theory and the newer one can be compared despite the fact that the central concept in the two theories is different —despite, that is, this lack of common reference (1989, p.429; italics in text).

This seems to be, according to Shapere, the adequacy requirement for a theory of scientific change. An adequate theory of scientific must allow comparability and ultimately judgement of such cases as the phlogiston versus oxygen theories, ether theories versus special relativity, caloric versus kinetic theories of heat, and, in general, all sorts of theories in which an old concept (and the corresponding entity) is abandoned and replaced, either by a new concept (and a new sort of entity or a new way of viewing a domain of investigation and research. What does this adequacy requirement indicate about the problem of theory-comparison? At least, one thing: there is far more involved in the comparability of theories than the mere continuity of the sort focused upon by reference theories, like Putnam's.15 Thus, in addition to the objections already raised against reference theories in Part II, there is the
failure to solve "one of the most important problems about theory-comparison." By focusing almost exclusively on the achievement of continuity of scientific discussion, through the allegation of common reference, they fail to confront (much less resolve) the problem of comparing scientific theories in which commonality of reference of the central terms in the theories is absent. "And this, as Shapere pointed out, is a, if not, the crucial possibility to be explained by any adequate theory of scientific change" (1989, p.430; italics added).16

In the end, Shapere claims that his "chain-of-reasoning" approach is capable of dealing successfully with cases such as the ones mentioned.17 And he concludes by mentioning three types of reasons which are often given when discussing cases of scientific change involving abandonment, modification, and replacement:

(i) The evidence which was alleged to imply the existence of the entity may (for a variety of types of reasons) be shown not to exist; (ii) the explanatory work done or claimed to be done by the entity may be shown to be done by a newly proposed one better (on the basis of a variety of possible types of reasons) than was done by the old entity; (iii) or the old theory may be shown not to be able to account for its domain of responsibility (either its quantitative or qualitative failures to account for old items may be shown not to be mere problems of incompleteness, as I have called them elsewhere [1985], but problems of incorrectness, or new items may be discovered for which it cannot account (1989, p.430).

Clearly what I have said so far about Shapere's view is only sketchy, and far from being satisfactory. A full
presentation would have to include it in the larger context and discuss at least some of the following questions:

(i) What constitutes a "reason" in science? What sorts of considerations count as reasons in science? How the considerations employed in the history of science have evolved to become what can justifiably be called "reasons"? In passing, what is Shapere's characterization of a concept? How does it pay lip service to his reason-based approach?

(ii) What constitutes a "background context" in science? Where does it come from? What is its role in experimentation, theory-formation, in short, what is its role in the process of meaning-making in science?

(iii) Finally, what are the implications of Shapere's view for philosophy of science, and for philosophy in general? Does it offer a true alternative? A viable and plausible alternative?

I will begin by considering what in Shapere's view constitutes the "background context" within which scientific research, both theoretical and experimental, is carried — at the concrete level of actual working science. If we consider, as he suggests, a sophisticated modern scientific experiment, like the still on-going solar neutrino experiment (Shapere, 1982), what constitutes the "background context" within which such an experiment is conducted? What sorts of considerations in other words actually guide the
conception, the performance and interpretation of the experiment? Clearly, the considerations which enable scientists to raise the problem, to determine the kind of evidence to seek and the methods by which to seek it, and much else, consist according to Shapere of a large variety of prior beliefs. Included among such prior beliefs in the case at hand are:

[T]he theory of stellar structure; the theory of stellar evolution; the theory of the sources of stellar energy; the theory of nuclear reactions; vast amounts of empirical data on nuclear reaction rates which cannot be derived from nuclear theory; weak interaction theory; knowledge of the properties of rare gases; technological capabilities of the instruments employed, both on a general level and more specifically regarding the idiosyncrasies of the particular instruments being employed in the experiment; such specifics as to how to clean the chlorine tank so as not to affect the experiment in potentially adverse ways; and so on (Shapere, 1989, p.431).

In Shapere's view, it is these prior beliefs which provide the background context in which the specific scientific research here question is conducted. He claims in fact that "without that body of background beliefs, or at least a good part of it, the solar neutrino experiment would have been, in a very literal sense, 'inconceivable'" (1989, p.431). Presumably, this example is typical of what happens in the formulation, execution, and interpretation of work in a wide range of cases of sophisticated scientific research. In general, however, one would tend to think that the network view or contextual theory is improved by restricting
rather than enlarging the amount of theoretical background that is (semantically) relevant. 18

Having said what constitutes a "background context," Shapere is also concerned with making clear what it is not -- in order to dispel in advance any possible misunderstanding. For example, the background context of beliefs which is brought to bear in a specific scientific research does not have, Shapere claims quite rightly, the kind of unity Kuhn's "paradigms" were supposed to have, or that a single "high-level background theory" is supposed to possess. But is it more like Kuhn's later notion of "disciplinary matrix" -- with several variable components? (cf: Part I, chapter 2). According to Shapere, it is much more of a hodge-podge, and many of its constituents cannot, in any reasonable sense, even be called "theories." In many cases furthermore, such items may not even be clearly consistent with one another. However, this is not to say that it is an arbitrary matter which background beliefs are brought to bear in a particular case. Though Shapere admits that there is often room for controversy about the relevance of certain elements of alleged background, he insists that it is by no means the case that 'anything goes'. 19 On the contrary, he claims, in the case of a great deal of background, there is little if any alternative; and in particular cases, one can explain why there is not.
Paradoxically, despite the fact that he has been critical of the historical approach pioneered by Kuhn and Feyerabend, no philosopher of science has adopted a more thoroughly historical approach to understanding science than Shapere.20 It is in this context that his often repeated slogan "We have learned how to learn and think about nature" (1989, p.434) must be understood. Thus, to the question "where does the background context of beliefs come from?" he would naturally respond that "they come from the past of science" (1989, p.432). From the past of science, we learn why certain beliefs and not others play the role of guiding context. From this, we learn in turn about the sort of considerations that come to count as "reasons" in sophisticated science.21 According to Shapere, two important aspects of what are used as "reasons" in science are:

- First, that they have previously been found to be successful in accounting for a body of information for which they were held responsible, and second, that they are coherent with other beliefs with which (subject to certain specifiable conditions that I must pass over here) they were expected to cohere (1989, p.432; italics added).

For Shapere, "reasons" (rather than "meanings") become the focal point of his approach; and they are defined in terms of (i) success and (ii) coherence. [Elsewhere (1984), he had included a third consideration (iii) freedom of doubt (i.e., specific doubt as opposed to universal philosophical doubt).22 But it is clearly too strong in that it need not be satisfied, and this may be why he did not keep it in his
1989 paper).

Interestingly enough, Shapere characterizes science as a process which has developed by "internalizing" what has come to be regarded (only as a contingent outcome of inquiry) as 'relevant reasons', and by progressively insulating itself from the world of "external" considerations.23 And he stresses repeatedly that it is still on-going and far from being completed. But does not such a consideration leave open the possibility, even the probability, of competing scientific communities who disagree about the comparative weights of reasons in specific cases and even about what constitutes a good reason in these cases. In other words, they may disagree about success, consistency, relevance and freedom of doubt. The internalization process becomes more and more domain-specific as science matures. Shapere recognizes this, but he is too often content to speak about 'science' as a whole, or as if it were a whole.24 There is nothing to prevent the internalization from resulting in variant forms or "cultures" of science within a given area of specialization (Shapere, 1984,p.322). For all that Shapere tells us about scientific reasons, he must allow that reasonable persons may differ. And for all his convincing arguments that objectivity and relevance are simultaneously possible, he does not offer us "closure mechanisms"
sufficient to explain why scientists actually agree as much as they do. It may be because such mechanisms cannot be formulated in general, and ought to be always relativized to a given domain, and approached on a case-by-case basis.

In any case, Shapere claims that many profound and pervasive changes resulted from this process of internalization: (i) successive reorganizations and redescriptions of subject matters for investigation (or "domains" as Shapere calls them (1984, p.xixii-xxiii, chap 13); (ii) rejection or modification of what had previously been considered to be background information; (iii) formulation and reformulation of criteria of qualification for a consideration to count as a "reason"; (iv) changes in what counts as "evidence"; (v) changes in what is to count as "observational"; (vi) changes in what a good "account" or "explanation" is supposed to do; and even (vii) changes in what counts as learning these things. According to Shapere, in all these cases, a context background of beliefs shapes, but does not strictly determine, the problems, methods, standards, and goals of science; and it is specific pieces of belief, functioning in specific ways, that do that shaping. He writes:

And in all such instances of change, when and to the extent that it became clear what were to count as internal considerations guiding inquiry, then, and to that extent, it became increasingly possible to see the process as having a self-corrective and evolving continuity of development (even when old entities were
abandoned), in which successive stages were linked by chain-of-reasoning connections (1989, p.433).

Among other changes, the scientific process often leads to rather drastic changes in our ways of thinking and talking about the world or our experience of the world. Early proponents of the contextualist theory of meaning express this idea by saying that the "meanings" of the terms used in science change, and that the reason why those "meanings" have changed is that they are determined by their theoretical context which changed. Shapere thinks that this way of describing scientific change is not only wrong, but has been the source of serious misunderstandings and misinterpretations of science. It has placed the focus on something presumably prior to science, namely philosophy of language, and distracted attention from the analysis of the way(s) in which scientific reasoning actually takes place.

While Putnam, for example, arrives at the conclusion that there are no aspects of the description of a thing or kind that are immune to revision on the basis of reflection on the nature of language, i.e., from a linguistic point of view, Shapere instead accepts the same conclusion on the basis of reflection on and acceptance of the openness of the scientific process to changes in light of reasons. And he claims that, as philosophers of science, we ought to be trying to understand those changes, and those reasons.

According to Shapere, Putnam was right in rejecting the
traditional accounts of meaning, but he failed to look at science and take into account actual scientific practices in order to deal with and explain scientific change. He assumed implicitly that if theory-comparison and commensurability is not accomplished through comparisons of meanings, then it must be done through commonality of reference. This assumption, Frege's legacy, and the idea that these two concepts "lie at the heart of philosophy" has in various ways dominated much of 20th century Anglo-American philosophy (see Part I, chap 1). Shapere seems at times to claim that this is part of the problem, and so, as we have seen above, he proposes a different approach, one which is neither a "meaning approach" nor a "reference approach" but a little bit of both, with a few distinctive marks. He writes:

The approach that I have taken, by making concepts depend on beliefs, rejects the view that there is a subject ("the philosophy of language", or whatever) which, independently of the content of science, can decide about what science is, does, and aims at. My view focuses centrally on reasons for beliefs, not on the way we talk about the world; for how we talk about the world, on my view, as well as what we are to make of that talk, is a product of science, something we have had to learn like anything else (1989, p.434-5; italics and underlining added).

Let us now raise explicitly the question mentioned at the beginning of this chapter, namely, can there be a viable and plausible alternative approach to the process of meaning-making in science, which is neither a "meaning approach" a la Kuhn and Feyerabend nor a "reference
approach" a la Putnam and other reference theorists? If Frege's legacy is part of the problem in trying to give an account of the process of meaning-making in science, then it behooves any philosopher of science to explore another alternative. If such an alternative exists, Shapere has well charted the course on which it can be found.25

Scientists have and give reasons for believing the "meaning" of a scientific term to be what they claim it to be at a given point in time; most importantly, they give reasons for changing the meaning of a term; if we trace of the chains-of-reasoning connections linking the various usages of the term, we would find, Shapere claims, continuity of meaning and commensurability. On his account, theory-comparison is not accomplished via comparison of meanings, nor is it accomplished by positing a commonality of reference, but rather by examining the CORC's linking the various stages of development of a given scientific concept, or the various usages of the term corresponding to it. But is Shapere's point in this respect convincing? Does the fact that scientists advance reasons in the process of changing meanings give us grounds for referential continuity or stability?26 It should be noted that the claim that reasons determine meanings is characteristically ambiguous, although textual evidence suggests that it is not the trivial idea that different reasons tend to lead to
different concepts. It appears instead to be the view that the chain-of-reasoning is itself actually part of the resulting concept—hence the necessity and not just desirability of the detailed historical investigations that Shapere urges us to undertake in the case of each scientific concept under consideration. But, if that is the case, does it not follow that two scientists could not arrive at the same concept by different routes? If this follows, then there would little or no merit to this particular proposal.

If Shapere shifts attention away from "meanings" to "reasons", he also redefines the nature of the relation between "meanings" (or rather reasons for beliefs) and "reference." Like earlier contextualists, Shapere admits that what is done in science, whether one talking about meanings or reasons, is shaped by a background context of prior beliefs. But unlike earlier contextualists, he does not believe this background context to be monolithic, or all-determining, as they have assumed; it is this belief principally which has led to incommensurability and relativism. Instead, for Shapere, the background context could be quite varied and have many disparate components; it only shapes what is done or said in science. "Even background beliefs can be rejected, and for reasons" (1989, p.435).

As we have seen, Shapere has attempted to show that a
proper understanding of the way contexts of beliefs function in shaping scientific research, in terms of the role of that context as specific reasons, can resolve the problem of theory-comparison. In his view, the central focus for philosophers of science should have always been scientific reasoning, as it actually takes place. This remark is not intended just to suggest that attempts to interpret science through the philosophy of language, logic, or whatever, have only succeeded in grossly distorting the scientific process; nor is it intended just to suggest that looking at science in its own terms will resolve some of the problems involved in scientific change. As Shapere writes:

The point is more general, more in keeping with a larger body of knowledge at which we have arrived. For philosophy of science, like any other discipline, must, in a truly genuine naturalistic spirit (comparison and contrast with Quine, Shapere, 1987), construct its interpretation of the knowledge-seeking enterprise in terms of the best background information available... Given that background context, the rational task of philosophy of science, and, of philosophy in general, must be to seek, within that large framework, a historical account of the development of our ways of thinking which, while it avoids the absolutism of approaches that pursue a higher and surer a priori basis for interpreting the scientific enterprise, also makes intelligible, without falling into a superficial and uncritical relativism, the process of reasoning that has led us toward such understanding (1989, pp.436-7; italics and comments added).28

He also reminds us in much of his work that the world-view which has emerged in the history of science has altered our expectations so profoundly that even what we must accept as an explanation, an account of our experience and the
universe in which it takes place (including our understanding of what something must be like in order to exist), and of what should count as a reason in favor of or against such an account, have been radically transformed, in ways that could not have been anticipated by any a priori considerations. According to Shapere,

Our past experience with the development of science suggests that new theories may turn out to be wholly unlike anything we have so far been able to imagine" (1984, p.338).

The emergence of that world-view, and the revisions it has engendered in what we can expect understanding to consist in, must now constitute the background context, the framework, within which we must seek an understanding of how it has emerged and of what its implications are. It is, in short, both the object of our inquiry and the context within which it must proceed. (1989, p.436). Indeed, one of the disturbing lessons that we have learned from the history of science is that virtually every major theory is bound to fail at some point or other and that our conceptual framework in the year 3000 for example, should we, as a species, survive so long, is likely to be quite different from our framework in 1990. "That is a sense of possibility that we must make room for in our philosophical theories" (1984,p.338).29

To many philosophers of science of formalist persuasion, Shapere's proposal may look like a contribution
to the "new fuzziness" in that it lacks closure mechanisms, i.e., crisp, logical, or methodological formulas and principles. Shapere's response to such a charge might well that it is not his fault if the world and the scientific enterprise are messy..."The single Truth, and the precision simply are not there..." (1984, p.220). In the end, I think that Shapere's proposal is very promising and deserves to be considered despite remaining difficulties and problems. As we shall see next, Nersessian's integration of some of its main insights attests to this.
Chapter 2: Nersessian’s Cognitive-Historical Approach

Nersessian, much like Shapere, is a proponent of the "new philosophy of science": she holds the major tenets discussed in the introduction of this project, and attempts in her work to exemplify them. She is not only interested in bringing out the deficiencies of past and present accounts of "meaning" and "meaning-change" for scientific theories,1 but more importantly in arguing for "the appropriate method by which to obtain an adequate account" of the process of meaning-making in science (1987, p.162). That is in fact her expressed intention in "A Cognitive-Historical Approach to Meaning in Scientific Theories" (1987, pp.161-177). Nersessian further illustrates the method she has in mind with two specific problems: (i) how to provide an adequate representation for the "meaning" of a scientific concept and (ii) how to analyze the role of imagery and analogy in concept formation and development in science. For the latter purposes, she considers the case of the "field" concept from Faraday to Einstein (1984).

Turning to Nersessian’s appropriate method for obtaining an adequate account of the process of meaning-making in science, it must be noted at the outset that she makes this methodological proposal in order "to fully incorporate the dimension of discovery—the history of science and the science of cognition—into the philosophical analysis of the conceptual dimension of
science" (1987, p.164). 2 For Nersessian, our understanding of science would remain seriously deficient if we fail to examine the question of how scientific concepts emerge and are subsequently altered. For, as she points out quite pertinently:

The creation of concepts through which to comprehend, structure, and communicate about physical phenomena constitutes much of the scientific enterprise. Concepts play a central role in the construction and testing of the laws and principles of a theory. The introduction of new concepts and/or alteration of existing ones is a crucial step in most changes of theory. And, in many scientific controversies what is at issue is disagreement over the interpretation of fundamental concepts. In short, articulating concepts is a central aspect of scientific research (1987, p.161).

According to Nersessian, two sets of problems must be addressed by an adequate account: (1) representational problems and (2) developmental problems. The former have to do with such questions as how to specify the constituents of the 'meaning' of a scientific concept and where to locate these concepts (in the head of scientists? in the scientific community? or in a Platonic realm?). The latter have to do with the nature of the processes through or by which concepts are formed and subsequently changed, the role and place of imageries and analogies, (and one should add metaphors and models) in concept formation in science.3 And Nersessian argues, "[t]he specification and solution of both representational and developmental problems require a method which is 'cognitive' and 'historical' (1987, p.163).
Why should the method be 'cognitive'? Because, Nersessian assumes rightly, I think, that "the cognitive mechanisms at work in the meaning-making dimension of science cannot be fundamentally different, i.e., different in kind, from those we employ in non-scientific and science-learning contexts" (Nersessian, 1988, pp.163-183). She also argues that "the problems encountered in characterizing the nature and process of concept formation and change in science can and should be examined in light of pertinent results, interpretations, and debates in the cognitive sciences, in particular, in cognitive psychology and in artificial intelligence, where it interfaces with psychology" (1987, p.164). Cognitive science has not yet evolved to the point where we could consider wholesale transfers of analyses of human cognitive mechanisms to the scientific case. It may not always be desirable to do so. One must recognize however with Nersessian that "any adequate science of cognition must also take the data from the analysis of science into account in its formulation, and this has not been done to any significant extent" (1987, p.164; see also Tweney, 1981;1985;1989; Giere, 1988; Thagard, 1988.4

Why should the method be thoroughly 'historical'? Because, in Nersessian's view, "this is one instance in which historical analysis is necessary for the philosophical problem" (1987, p.163). The problems of an account of the
process of meaning-making in science cannot be formulated, let alone resolved, merely by comparing the developed concepts of various scientific theories [see Part I, chap.3; see also Shapere (1989) and Kuhn (1988) for a similar point]. For example, she claims that we cannot simply compare Newtonian 'mass' and Einsteinian 'mass'; what we need instead is a detailed examination of how we got from one to the other. i.e., of all the intermediate steps. She later makes a similar point about Faraday's concept of 'field' and Einstein's concept of 'field'. If the early positivist accounts were decidedly ahistorical, Nersessian argues, the accounts (of Kuhn, Feyerabend and others presumably which have claimed to establish incommensurability as a fact of history are decidedly unhistorical.5 And she concludes, like Shapere:

What is lacking from them are fine-structure analyses of the period of transition between theories. Where these have been provided, the 'problem of incommensurability of meaning' appears not all that profound (1987, p.163; italics added).

Concerning the first example, she points out that we did not go directly from Newtonian mechanics to Einstein's special relativity, but passed at least through the theory of electrons (1987). Concerning the second example, she shows that similarly we did not go from Faraday's 'field' to Einstein's, but passed at least through Maxwell's 'field' and Lorentz's 'field' (1984). According to Nersessian,
What examination of the history of science shows (...) is that the real problem is one of accounting for continuous development which is not simply cumulative, i.e., of accounting for 'meaning variance' with commensurability (1987, p.163).

Indeed, from Nersessian's point of view, even if there may be cases of "localized" incommensurability, there is a far greater degree of commensurability in the scientific enterprise. And so the task, I take it, is to provide an account of the various kinds of meaning-variance that may occur within the larger background of commensurability. I think that Nersessian's point in this respect is well taken and rather plausible.

Nersessian further characterizes the distinctive historical dimension of the method she recommends by arguing for a different approach to case-studies of particular episodes in the history of science. She rejects the kind of case-studies which have been undertaken by many post-positivistic philosophers of science only to serve as tests, illustrations for, or counterexamples to particular philosophical theories. In most of these cases, she argues, the historical data have been distorted to fit the needs of a particular philosophical theory, or to make a particular philosophical point. In an attempt to formulate one of the main tenets of the "new philosophy of science", she writes:

The history of science needs to be examined with a philosophical sensitivity to the problems at hand, but not with the goal of fitting the data to presupposed
solutions. We should be open to the possibility that some of our problems are not real and have a willingness to abandon proposed solutions as well as problems. It is, of course, not possible to come to the history 'pure', but as far as possible we should be cognizant of our preconceptions and try not to adapt the data to these. Our proposals should grow out of the historical data and should not be imposed upon them. Rather than only the 'historian's task', the 'spelling out of how particular concepts developed with time' is the starting --and end-- point of a philosophical analysis of meaning in scientific theories (1987, p.163; italics added).

In response to the common objection raised against such a 'historicist' perspective in philosophy of science, namely, that the 'philosophical reward' of such analysis is slight and slim, and even irrelevant (Enfield, 1985, p.641; Leplin, 1988), Nersessian (and Shapere) would contend instead that the rewards are not just relevant but may in fact be great, when such a perspective is adopted. But does that mean that the philosopher must become a historian as well --at least for a time? Probably. Does that mean that philosophy of science must forego its normative concerns --and focus exclusively on descriptive considerations? Certainly not. But it does mean that the philosopher's normative concerns must reflect or be based upon descriptive considerations, or else not be in flagrant contradiction with these considerations (cf: Kuhn, 1988). 7

Before considering more explicitly Nersessian's solutions to both the representational and developmental problems, it is important to note two fundamental claims that she arrives at as a result of her preliminary
historical analysis of the concept of 'field' (1987, p.166): (i) The creation of a scientific concept takes place within networks of beliefs and in response to specific problems. These beliefs and problems are theoretical, experimental, methodological, and metaphysical in nature. In Nersessian's view, we cannot separate questions of meaning from questions about beliefs and problems. For, she claims, when we examine the "chain of reasoning" connections leading from one phase of development to the next, we can see that changes in meaning(s) come about in response to changes in beliefs and problems --and vice-versa, as Kuhn (1988) points out. These changing networks of beliefs and problems which are in part communal and in part individual, overlap in such a way as to provide continuity and diverge enough to create fundamental changes. Nersessian's point is that if we are to understand 'meaning' and 'meaning-change' in science, it is necessary to include these networks in our analysis. But the following questions arise: Is Nersessian thus enlarging or restricting the amount of background knowledge that is semantically relevant? Is she making it by the same token more holistically determinant? Unlike previous CTM, she is clearly enlarging it and somehow specifying it; however, like previous CTM, she seems to view it as holistically determinant. In general, one would think that a network view like Nersessian's is improved by restricting rather
than enlarging the amount of background knowledge that is semantically relevant, and by rendering it less holistically determinant. For, by virtue of the network-metaphor sustaining this view, a change in some aspect or other of the network entails changes in other aspects of the network; and the larger and the more determinant the network, the more significant and radical the conceptual change will be. Does not this raise once again the specter of incommensurability? I will reconsider this possible objection in a moment.

(ii) In each theory, the meaning of a given concept may differ significantly. Yet, each concept shares in part of the meaning of its predecessors — and it shares more with its immediate predecessor than with those more remote. Thus, while there is may be meaning-variance, there is a significant degree of continuity and commensurability; and this continuity and commensurability can be established by tracing the chain-of-reasoning connections which have led from one concept to the next, and so on.

One may at this point object to Nersessian, and argue that her view is in fact similar to the "contextual theory of meaning" which has given rise to incommensurability (Enfield, 1985, p.642). Nersessian acknowledges that her view is a contextual or network view, but she claims that it is one which somehow escapes the incommensurability problems
inherent in other contextual or network views through the device suggested by Shapere, namely, the "chain-of-reasoning" connections [COCR's]. In both Nersesian's and Shapere's view, a concept ends up being a set or rather a cluster of reason-related criteria, and the problem becomes one of representation of the concept in question which encompasses these criteria. Furthermore Nersesian claims that "referential stability is not a problem because continuity of reference is built into the synchronic and diachronic representation of the concept" (1987, p.176, note 6). We shall now see more precisely whether and how this is so, and turn for this purpose to the representational problem. In anticipation, it should be pointed out that though she seems concerned with referential stability, she does not give us any clear and convincing means to comprehend how and why it is not a problem. In fact, like Shapere, she does not have much to say about 'reference'; instead, she concentrates on 'meanings'—how to characterize them at some point in time, and over time. In contrast, the model proposed by the new cluster theory that I envisage (Part IV), will be concerned with both 'meanings' and 'reference', at some point in time and over time, as they may both change. Its main point in this respect is that we should neither postulate referential stability nor seek to establish commonality of reference at
all costs.8

In Nersessian's view,

The most pressing representational problem is that of specifying what constitutes 'the meaning' of a scientific concept. A prerequisite to any characterization of how concepts form and change is a determination of what is to be included in the meaning of a concept. To allow for an account of 'meaning-change', the representation of a scientific concept must be both synchronic and diachronic. That is, the representation must stipulate what constitutes 'the meaning' of a concept at a point in history and as it changes over time, either intra- or inter-theoretically (1987, p.166; italics added).9

Following her recommendation, the synchronic representation of a concept must then be determined first. She considers the 'definitional' representation of a concept, in terms of a set of necessary and sufficient conditions, and rejects it as inadequate in part for the reasons discussed earlier in Part I, chapter 3. In passing, I think that she mistakenly believes that all previous philosophical semantics with which she contrasts her view 'defined' a term in terms of necessary and sufficient conditions. In fact, as suggested, there have been some variations on this 'definitional' representation of a concept. In any case, she finds that such concept-representation does not do justice to the historical data which she was able to gather in her case-studies [e.g., on the concept of "electromagnetic field" (1984; 1989); that of "ether" (1984)].

Concerning the 'field' concept, she writes:

Given its central position in modern physical theory, we can learn much about the conceptual dimension of science
by examining its development and its role in the formulation of several theories (1987, p.165).

If the data in question is indeed representative of much of concept formation and development in science, as Nersessian thinks it is, then the arguments against a 'definitional' concept-representation are seriously damaging and her proposal is worth being considered. She says in this respect that "although the details of a satisfactory representation have yet to be worked out," she will present "a promising avenue of resolution for this representational problem (1987, p.167).

As expected, she begins by distinguishing between concept-type, the general concept of 'electromagnetic field' for example, and concept-tokens, instances of the general concept, such as the electromagnetic field concept of the theory of electrons. Then the meaning of each instance, or concept-token, is represented synchronically, at some point in time, by a multi-component 'vector'. The salient components of the vector must be determined in each case, and these may be different for different kinds of concepts, such as substantive concepts and dispositional concepts. According to Nersessian, for a quantitative substantive concept such as 'field', some components would include its ontological status and reference, its function or behavior, its causal properties, and its structure (mathematical or otherwise).
To obtain a diachronic representation, Nersessian proposes to expand the vector into an 'array'—what she has elsewhere (1984) called a 'meaning-schema'—which exhibits how the various components changed or did not change over time. It is worth noting that Nersessian's notions of 'vector' and 'array or meaning-schema' seem to have much in common with the notions of 'schemas', 'frames', and 'slots' developed in cognitive science (i.e., AI). This may be an indication of the fruitfulness of the "cognitive-historical approach" advocated by Nersessian. Perhaps, the concept-representation, which she is proposing and which presumably grew out of the historical data, can even be developed further in light of other notions formed in the cognitive sciences (cf: Thagard, 1988; 1990).11

In Nersessian's view, what has emerged from recent discussions about concept-representation (e.g., by Rosch, 1975) and from renewed interest in Wittgenstein (1953, #65-88) is that "it is possible to represent concept in such a way that while they are indefinable, they are complex—although there is far from a consensus on how to do this" (1987, p.168). She considers two predominant conceptions: (i) the prototype view and (ii) the probabilistic view (see Smith and Medin, 1981). In both views, a concept is represented as a set of weighted features. However, as Nersessian explains:

On the 'probabilistic' view, this set of features is
fundamental, and a prototypical instance, i.e., one which maximizes the weighted score, is generated as a by-product. With the 'prototype' view, the prototypical instance is itself basic, and it determines the weight of the various features. (1987, p.168).

What is important here with these views is that they both represent a concept in such a way that the features or properties which serve to characterize its 'meaning' have a substantial probability of occurring, and are not considered to be 'essential', or 'necessary and sufficient'. Instead, they may be considered as a cluster of properties. And the cluster of properties, or as Nersessian puts it, "[t]he overlapping set of resemblances" makes the concept into a unit" and entitles us to individuate it as the concept 'X' or 'Y'. Furthermore, it allows for the possibility that the representation can change over time.

In sum, a concept can be represented at some point in time and over time --which is what the nature (open-ended) of concept formation in science in fact requires. The representation of tokens of scientific concepts is the set (or rather cluster) of salient components of its 'vector', and that of the type, a set (or Cluster) of family resemblances exhibited in its 'array' or 'meaning schema'. However the following questions must be asked: How do we determine the components of the vector? of the meaning-schema? How do we select the salient ones? What is a 'resemblance'? How do we weigh the relative importance of resemblances? How do we assign relative weights?
Despite these remaining questions which could be the source of serious problems, Nersessian argues that

Of the two views (probabilistic and prototype), the 'probabilistic' seems more in accord with the scientific case. The set of features is fundamental, and a prototype would be generated from the 'meaning schema'...[However] the main point that I wish to make here is that either view is better suited to our purpose than the 'definitional' view. They both offer the potential to represent development, continuity, and change in a way the definitional notion cannot (1987, p.168).13

About the 'field' concept example, she claims that we would be able to attribute a field concept to Faraday quite early without having to attribute either all the features of his mature concept or those of the modern conception. We would thus be able to show that its specific features developed over time and that it has features quite unlike other field concepts; and yet still be able to maintain that it is connected with other field concepts, in particular the modern one [see Table 1 below, adapted from Nersessian, 1984,p.158; 1989, p.328]. According to Nersessian,

In this way the 'problem of incommensurability of meaning' is overcome: the earlier and later forms of a concept bear a familial relationship to one another, and that is sufficient (1987, p.168; italics added).

But is it truly sufficient? --as she concludes using Shapere's line of argument and his notion of 'chain-of-reasoning' connections [CERC's].14

Table 1: Concept-type: 'Electromagnetic field'
<table>
<thead>
<tr>
<th>Ontological Status</th>
<th>Function</th>
<th>Mathematical Structure</th>
<th>Causal Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>F Substance (preferred) or state of aether</td>
<td>Transmits electric &amp; magnetic actions continuously through region surrounding bodies and charges</td>
<td>Unknown</td>
<td>Electric Magnetic effects</td>
</tr>
<tr>
<td>M State of mechanical aether</td>
<td>Same/ plus optical</td>
<td>Maxwell's equations</td>
<td>Electric magnetic optical effects, radiant heat, etc.</td>
</tr>
<tr>
<td>L State of non-mechanical aether</td>
<td>Same</td>
<td>Same/plus Lorentz force</td>
<td>Same</td>
</tr>
<tr>
<td>E State of space (ontologically on a par with matter)</td>
<td>Same</td>
<td>Same/plus Relativity interpretation</td>
<td>Same</td>
</tr>
</tbody>
</table>

F= Faraday; M= Maxwell; L= Lorentz; E= Einstein. The dotted lines indicate "chain-of-reasoning' connections [CORC's]'/.

Although Nersessian recognizes that there were embryonic field concept before Faraday, she begins her analysis with his conception because it is in his work that "the notion that an understanding of processes in the region surrounding bodies and charges is required for the description of electric and magnetic actions first took hold" (1987, p.165).15

As I mentioned at the beginning of this discussion,
Nersessian is not only interested in presenting the appropriate method by which to obtain an adequate account of the process of meaning-making in science, but also in illustrating the method in question with two problems: the representational problem and the developmental problem. Having just shown how to adequately represent the 'meaning' of a concept at some point in time and over time, she then turns to the developmental problem for a more detailed discussion. Her intention is to consider some of the processes by which scientific concepts are formed and subsequently altered. She writes quite rightly:

Contrary to those who have claimed that new concepts arise by some inexplicable 'creative leap' that defies rational analysis, when we examine the reasoning leading to the emergence and alteration of concepts, it is possible to discern methods of concept formation which are subject to rational analysis. Study of these methods shows that the dynamics of (conceptual) change are not those of abrupt, cataclysmic revolutions. However, such analysis does not yield Baconian 'sausage-making machines' for concepts either. Rather, what we have are certain recurrent strategies for the articulation of scientific concepts (1987, p.169).

To consider such strategies does not 'negate' the 'creative' nature of scientific concept formation, since, development, refinement, and application of these strategies testifies to the creativity involved. Furthermore, such a focus provides support and substance to the long-standing claims by some cognitive psychologists about concept formation: namely, that new concepts are also formed by 'accretions', 'deletions', 'combinations', 'differentiations', and
'shifts' from existing concepts (see Rumelhart, 1980; see also Goodman, 1978; Thagard, 1984).

So, instead of Gestalt theories (like Kuhn’s) —which have presumably emphasized radical conceptual change, whereby new concepts replace old ones as a result of scientific revolutions, we need to appeal to Progressive theories, which characterize conceptual change in terms of additions, deletions, combinations, reconfigurations, differentiations, etc. Contrary to Thagard (1990, pp.194-5), I think it is more appropriate to call the latter theories "Progressive" rather than "Accretion theories," since conceptual change may involve not just additions or accretions (of properties or relations), but deletions and reconfigurations. Also, contrary to Nersessian's view, I think that we need Progressive theories in addition to Gestalt theories, for we want to be able to account for different kinds of conceptual changes in science, i.e., radical and minor changes (see Part IV for further details).

According to Nersessian, what is lacking from Gestalt theories is an account of how or by what processes concepts are combined, differentiated, etc. Thus, she proposes to focus on the use of imagery (i.e., pictorial representations) and analogy made by Faraday and Maxwell for example in the construction of their respective 'field' concepts (see also Miller, 1984; Wise, 1979; Tweney, 1985).
As I pointed out earlier, a more complete study of the processes here in question would have to include an analysis of the role of metaphors and models as well [Part IV].

What makes Nersessian’s discussion of the role of imagery and analogy in the construction of Faraday’s and Maxwell’s ‘field’ concept interesting is that the issues involved can be extrapolated to the general problem of concept formation and development in science. Also, in Nersessian’s view, a good deal of recent work in cognitive psychology can be brought to bear on these issues (see also Thagard, 1990). For example, she claims that analogies, in their assimilating and articulating functions have much in common with the "structure-mapping theory" of analogies developed by Dedre Gentner and her colleagues (1983). Initially, this theory was designed to account for empirical studies of how people use analogies in learning new scientific concepts. Although there is clearly a difference between this cognitive activity (learning new scientific concepts) and the cognitive activity which consists in creating or formulating scientific concepts, they are related to each other in a significant way. We might be able to show how if we could compare the use of analogy and imagery in the articulation of concepts with the claims of the structure-mapping theory (see Nersessian’s "Conceptual Change in Science and Science Education," 1988, pp.163-183).
Nersessian is fully aware that the structure-mapping theory as it is presently formulated is not totally adequate for even the case-study of the 'field' concept, but she recognizes that when combined with the historical data, it offers a promise of deeper understanding of the roles of imagery and analogy in scientific concept formation and learning.

According to Gentner's structure-mapping theory, scientific analogies are systems of connected knowledge, and in its meaning-giving aspect, evaluation of an analogy is considered in terms of its explanatory power and not its validity as an argument. And as Nersessian explains:

"The theory assumes that the predicates of the base domain are brought across to the target domain as identical matches and not similarities. It assumes that 'similarity' can be construed as an identity between some number of component predicates. The difference between an analogy (a 'non-literal similarity') and what is called a 'literal similarity' is that the predicates that are mapped in an analogy are mostly relationships and not attributes, whereas with a literal similarity many more attributes are brought across. A simple example of a literal similarity is one between our solar system and one in another galaxy; of an analogy, one between a solar system and an atom (1987, p.174).

Nersessian does not offer (yet) a detailed analysis of analogies as 'structured maps' but she provides some support for the structure-mapping theory as a potentially fruitful theory for understanding the roles of imagery and analogy in scientific concept formation. She shows for example that the analogies used by Maxwell are truly systems of connected..."
knowledge, and that they are not used as arguments, but simply as heuristic devices for exploring certain representational possibilities. In that sense, they exemplify the type of analogy examined by the structure-mapping theory. She writes:

In the process of articulating the field concept, certain features of the domain of the analogue were incorporated into it. They were primarily relationships. For example, a current does not have the attributes of the little spherical balls used to represent it; rather, the dynamical relationships between the balls and the vorticities are those assumed to hold between an electrical current and magnetism. One way to interpret Maxwell’s repeated claim that no physical hypothesis is being made is as a claim that attributes are not being mapped from the domain of the analogue. Also, the structure-mapping theory holds that analogies mapping systems of relations (2nd-order or higher relations), such as causal relations, have higher explanatory value. This is borne out in the Maxwell case, where the analogy (...) went further in the articulation of the field concept by assuming identities in causal relationships between the two domains (1987, pp.174-5).

Finally, Nersessian points out that the continuum proposed by the theory here in question between literal similarity (mapping mostly attributes) and an analogy (mapping mostly relations) makes it possible to consider the roles of imagery and analogy on a continuum as well. As Nersessian has shown in her historical analysis, many of the attributes of the image of the lines of forces are mapped into Faraday’s concept of field. However, with the images used to represent Maxwell’s physical analogies, it is the relationships represented by the image, rather than the
attributes, which get mapped. Thus, the function of the imagery has somehow evolved from concrete to abstract. When they function concretely, images can be considered to be 'literal similarities', and when they function abstractly, they can be considered to be 'analogies'. As a literal similarity, the specific image is crucial, since many of the attributes it represents are incorporated into the concept. But as an analogy, it should be possible to use any image which represents the relevant relationships, since only the relationships are of interest. In the end, the changes in the function of the imagery employed during an episode of the history of a given science may well indicate the transition from a qualitative formulation of a concept to a quantitative one (e.g., from Faraday's to Maxwell's 'field' concept).

Some important questions concerning scientific concept formation have been dealt with more or less satisfactorily. Others however remain unanswered by both Nersessian's account and Gentner's structure-mapping theory. For example, how can an analogy be altered to fit or satisfy the needs of the domain under investigation? [Dynamical aspect of the use of analogy?]. How does the selection of an analogy take place? [Role of beliefs and problems in the selection process?]. In the end, perhaps the most significant testimony to the plausibility and viability of
Nersessian's account lies in the fact that answers to these and other questions along these lines may require the further application of the "cognitive-historical method." Perhaps, they will require instead that Nersessian's account be revised at least in part or supplemented by other kinds of considerations from other fields. Be it as it may, Nersessian's account deserves to be explored further.

In Part IV, I will attempt to show how the new and revised cluster theory incorporates the main insights of both Nersessian's and Shapere's account, insofar as the former continues and develops the research program inaugurated by the latter. I will also show more systematically in what respects the former differs and departs from the latter. This should be a straightforward task since, as I have suggested above, Nersessian and Shapere are in fact "closet cluster theorists" who deny that they are, because they only have in mind a traditional version of cluster theory. Leplin (1987; 1988), for example, attributes such a view to both of them --only to undermine it in favor of a revised version of Putnam's CTR. It thus appropriate to examine briefly Leplin's main argument against both Shapere and Nersessian before going further to consider Kuhn's most recent proposal (1988).

The philosophical problem that Shapere's and Nersessian's proposal is supposed to solve is
incommensurability of meaning. As meaning changes with theory, uses of a term in different theories do not express the same concept. Accordingly, theories cannot sustain the logical relations presupposed by attempts at comparative evaluation. The solution they seem to propose is that uses of a term in different theories are commensurable stages in the growth of a single concept if they are connected by COR's (chains-of-reasoning); if there are scientific reasons for the changes of theory reflected in differing uses of the term. That is supposedly the big breakthrough. In other words, the big breakthrough is that we look at actual science and discern rationality in its transitions, whereas the philosophical mistake which has given us incommensurability (and which has not been avoided by Putnam's CTR) was to focus almost exclusively on language. Now, according to Leplin, this "breakthrough" is by now second nature to a generation of philosophers; it is at most a methodology, not a philosophical theory, and it solves no philosophical problems. In fact, he claims, in Nersessian's application of the methodology, it exacerbates problems.

I think that Leplin's comments are themselves exacerbated. First, given that the "historicism turn" has been articulated in the early 60's, this "breakthrough" must be by now second nature for most philosophers of science, but I don't think that it is yet. Many philosophers of
science (including Shapere, Nersessian, Kuhn and myself) are still trying to draw the proper significance of this "historical turn" methodologically and theoretically speaking. Second, if what Shapere and Nersessian are proposing is at best a methodology, it is a methodology which aims at formulating an adequate philosophical theory of the process of meaning-making in science. Third, their proposal may not solve in a definitive manner any philosophical problems, including the problem of incommensurability, but it does offer, as I have shown in previous chapters, some new insights into some old problems, new ways of formulating these problems, and perhaps a better chance to solve them or to improve our understanding of these problems.18

In further objections to Shapere and Nersessian, Leplin makes the following points:
(i) Their methodology is based on an article of faith: every substantive change in theory has an autonomously scientific rationale, that chain-of-reasoning connections (CORC's) connect, albeit indirectly, everything that happens with everything else that happens in science.
(ii) The mere presence of CORC's can hardly serve to distinguish conceptual growth from conceptual diversity.
(iii) If incommensurabilists err on the side of diversity (holding at the extreme that every individual scientist is
somehow insulated from every other), Nersessian and Shapere err on the side of homogeneity (in that they suggest what amounts to a sweeping, indiscriminate monism).

Again, I think that Leplin is wrong on these three points. Concerning the first one, there is nothing in Shapere or in Nersessian that may lead one to think that they believe what Leplin claims. However, given that science is the process of inquiry that it is, reasons and evidence are provided perhaps more than in any other discipline for whatever changes are introduced, choices made, and decisions taken. As for Leplin’s comment on CORC’s, it is unfair because that is not clearly what Shapere or Nersessian intended for this notion. Only that there are CORC’s linking the various phases in the development of a scientific concept. And that as we trace the changes in the problems, beliefs, and interests responsible for the differing approaches and conceptions of the theoreticians involved, we may be able to better understand the process of meaning-making in science.

Concerning the second (and third) point, I find on the contrary that Shapere’s and Nersessian’s methodology does distinguish conceptual growth from conceptual diversity, but it seems rather optimistic about conceptual growth and perhaps too lax on conceptual diversity. And so, rather than erring on the side of homogeneity, as Leplin claims, it might err instead on the side of diversity —though not in
the sense defined above, but in the sense of being too reliant upon a case-by-case approach, and not offering enough "general criteria" or even "closure mechanisms."

This being said, Shapere/Nersessian's methodology leads to one important claim, namely, that the theoreticians responsible for the different stages in the growth of a single concept (e.g., 'field', 'electron') were all reasoning about the same problem. (Nersessian, 1984, p.155; Shapere, 1989, p.428). This in turn lead us to believe that concepts are somehow individuated by problems. But we are given no clear criterion for the individuation of problems or concepts and none for deciding whether or not the conceptions produced by reasoning about the same problem are phases in the growth of a single concept. The CORC device certainly does not constitute a theory of the process of meaning-making in science, but it seems to place the emphasis where it should, i.e., on process,19 and it can serve well to assuage worries about incommensurability in specific cases.
Chapter 3: Kuhn Revisited --1982-1990

Ever since Kuhn produced his famous work [SSR] in 1962 (supplemented in 1970 by a postscript), many questions and objections have been raised against his account of the scientific enterprise, and in particular of the process of meaning-making in science. Ever since however Kuhn has attempted to answer these questions and objections as best as he could. Still, in that process, he has not produced a fully developed theory of meaning. In recent times (1982-1990), Kuhn has returned to a set of themes (e.g., the role and place of history in philosophy of science, incommensurability of meaning) that he raised some 20-30 years ago, in the hope that he will be able to make further clarifications and lay out at least programmatically the outlines of the (kind of) theory of meaning that he thinks would be adequate. In this chapter, I shall discuss Kuhn's more recent contributions, with particular attention being paid to the theory of meaning that emerges from them. In anticipation, this theory appears to be a "locally holistic verification theory of meaning" based among other things on notions such as lexical structure, hierarchies of taxonomies, and feature spaces.1 Also, Kuhn argues explicitly that this theory is not a cluster theory of meaning. But, as I shall show, he denies that his theory of meaning is a cluster theory, because like other philosophers with the same predicament, because he has in mind a
traditional version of cluster theory. This discussion shall serve to situate the recent Kuhn (1962-1990) with respect to the early Kuhn (1962-1970) but with also respect to recent contributions (e.g., Shapere, Nersessian) on the subject of scientific development and scientific change, and most importantly with respect to the cluster theory that I envisage.

Originally, Kuhn was interested in "the nature and authority of scientific knowledge" and he approached it with the conviction that the long dominant views on this subject might be radically subverted if we paid closer attention to "what scientists actually do" (SSR, 1962, chap. 1). That is also what many "historicist" philosophers after him believed and still do (e.g., Shapere, Nersessian). The data about the behavior of scientists that he made use of was derived from various sources: personal experience, an emerging sociology of science, and mainly from historical case-studies of science. Looking back however, Kuhn thinks that he and others were in fact "misled by seeing history as a source primarily of data." According to the new Kuhn, the historical case-studies that he and others had undertaken, provided not only data but a perspective from which to view them. Though that perspective informed his data, he was apparently not fully aware of its role in his own work. However since, he has become convinced that "scientific
knowledge can properly be understood only as a product of history," of a continuous developmental process" (Kuhn, 1986; compare with Shapere and Nersessian).2 That is why Kuhn focuses on a set of problems concerning the nature and consequences of conceptual change. Though much discussed in recent years, these problems look different, according to Kuhn, when viewed as consequences of the nature of history rather than of the facts history provides.

To highlight that difference, Kuhn has argued that from the viewpoint of the historian, all knowledge of nature emerges from prior knowledge, usually by extending, but sometimes by partially replacing it. That generalization is as relevant to the so-called context of justification as it is to the context of discovery. To discovery, the prior body of knowledge supplies the conceptual tools, the manipulative techniques, and much of the empirical data required for the emergence of cognitive novelty. To justification, the same prior body of knowledge provides the only standard of comparison by which a candidate to succeed it can be judged. In the sciences, that is, the foundation for future knowledge is present knowledge, and there is no other foundation --more neutral, less contingently situated --to be had. Contributing to knowledge and evaluating contributions made by others are historically and culturally situated activities: no individual can engage in
either until he or she has mastered both the language of the community to whom the contribution is offered and also a number of that community's currently accepted truths.

This line of reasoning is clearly reminiscent of some of Shapere's claims and Nersessian's assumptions. As a descriptive statement about the way science actually develops, what I have just attributed to Kuhn about historicity can sound trivial, but he takes its import to be more than factual; as "somehow deeply implicated in the nature of knowledge itself" (Kuhn, 1986; cf. Shapere, 1984 for a similar point). If Kuhn is right that the foundation, cognitively speaking, of the science of one time is the science of the immediately preceding time, then, according to him, two distinct tasks are involved in providing examples for philosophers of science to analyze.

First, historians must somehow regain the past from which their narrative sets out; they must re-construct an older cluster of knowledge claims and re-establish somehow the nature of its appeal to those who held it. In this task, Kuhn argues that historians must behave like ethnographers striving to describe and understand the behavior (linguistic and otherwise) of an alien culture. Second, in their narratives, historians must show how we went from an older cluster of knowledge claims to an
expanded or revised successor. Because the "ethnographical aspect" of the history that he has in mind is less widely recognized than the narrative aspect, Kuhn attempts to bring it out more explicitly.

Kuhn introduces the problems presented by what he calls the "ethnographic aspect of history" --which precedes its narrative successor-- by discussing three examples from the history of science: (i) the concept of motion from Aristotle to Newton, (ii) the concept of battery from Volta to Faraday, and finally (iii) the concept of black body from Max Planck on. These case-studies reveal the past of science as alien (Kuhn, 1987, pp.7-22). They are supposed to provide concrete illustrations of both the need and results of that quasi-ethnographic task. Their central and shared feature is that each centrally involved a description, not only of the beliefs held by the community members, but also of the meanings of some of the words in which those beliefs were expressed. Of those terms, some had since vanished from use, while others, though still in use, now functioned differently, carried different meanings. Kuhn claims that what these examples have shown is that until the older meanings were recovered, many passages in the texts which recorded an older body of knowledge seemed nonsense. Furthermore, they make clear that in order to make those meanings comprehensible the study of meanings as
well as beliefs were required (see Shapere and Nersessian, for a similar point).

As mentioned earlier, historians, according to Kuhn, who study the development of knowledge of nature have a double task: They must provide a credible explanatory narrative concerning changes in ideas about an interrelated set of natural phenomena. But before they can provide such narratives of development which opens at some point or other in time, they must somehow "set up the stage," reconstruct the background context, by telling what people then believed. Mere quotations will not do. For Kuhn historians must instead approach older ideas in an ethnographic mode, one that aims to account for the coherence and plausibility of those ideas to the people who held them. Only if one understands why an older set of beliefs was held, what appeared to be evidence for it, can one hope to recount, analyze, or evaluate the process by which it was given up and replaced (Kuhn, 1982; 1990). One can hardly disagree with such a claim. In order to show the philosophical as well as the historiographic importance of this point, Kuhn explains that in each of the examples that he discussed, he described a set of past beliefs about some aspect of nature, and that, in order to do so, he had to describe in each case the meanings of a few of the terms in which those beliefs were stated. He remarked that these terms are not just any
terms; they are of a special sort; typically, they are what he calls "the names of (fundamental) taxonomic categories available to members of the speech community that use them" (1982; 1989; 1990). And as such, "they carry the community's ontology, supplying names for things which its world can and cannot contain." They function in Goodman's words as projectible terms, which can appear in laws of nature, counterfactuals, or in inductive generalizations; in short, they have the characteristics of "natural kind terms" (cf: J.S. Mill's description).

Kuhn stresses one point, and that is, even if for the purpose of exposition Kuhn artificially distinguished between descriptions of meanings and descriptions of beliefs, throughout his interpretive efforts, "beliefs and meanings were encountered together, in an inextricable mix." In order to understand "puzzling statements of beliefs," he had to engage in "rediagnoses of meanings," and in order to carry the latter, he had to take clues from the former. That is why, he notes quite rightly I think, that exercises of this sort are sometimes described as "attempts to break into the hermeneutic circle" (1982).4

In a second move, Kuhn reconsiders the issue once described as "incommensurability" and asks anew about "what makes the past foreign and about the extent to which and the manner in which its foreignness can be transcended."
The context in which he does so however is the following. He argues that the entanglements of beliefs and meanings are intrinsic to the nature of knowledge, for, part of what counts as knowledge at any given time is acquired during the process of learning the language in which that knowledge is stated. It is fair to say that in his view essential parts of a community’s knowledge of nature are embodied in the structure of the lexicon which members of that community share. To acquire a lexicon containing such Aristotelian terms as ‘motion’, ‘place’, and ‘matter’ is to learn things about the world: that a falling stone is like a growing oak, that nature can no more exhibit a vacuum or an inertial motion than it can a square circle. To describe Volta’s battery with the lexicon of electrostatics (no other was then applicable to the phenomena involved) is to make the battery like the Leyden jar, to locate the source of current at the metallic interface, and to specify the direction of the current. To employ the terms ‘resonator’ and ‘element’ as Planck did is to represent black-body radiation as like acoustic radiation and the energy element & as a sub-division in a continuum rather than as a separate energy atom. In each of these cases, the description of phenomena required commitment to a lexicon and that lexicon brought with it restrictions on what those phenomena could and could not be. If nature were later found to violate those
restrictions as occurred in each of the examples presented— the lexicon itself would be threatened. Elimination of the threat required not simply the substitution of new beliefs for old but alteration in the lexicon with which the prior beliefs were stated (1987, pp.18-19ff). Now, it is because lexical changes of the sort mentioned by Kuhn separate us from the past, even the quite recent past, in some areas, that past science cannot (always) be fully recaptured with our current lexicon. And he claims: once this is acknowledged, and once we have a model of how a lexicon works, then one can suggest meaningfully "how the past is to be recaptured."

In considering the changes of meaning which initially separate the historian from the past, Kuhn is in fact reconsidering an aspect of conceptual development which 25-30 years ago, he called 'incommensurability.' At that time, its primary application for Kuhn was to the relation between successive scientific theories. In that application, changes in word-meaning explained the characteristic difficulties of communication between proponents of competing theories. Corresponding conceptual changes were the basis for his talk of the Gestalt switches which accompanied theory-change. Looking back, that viewpoint still seems correct in its essentials to Kuhn, but in need of considerable modification of detail (1982).
According to Kuhn, there is one aspect of this modification that is particularly relevant here. In 1962, he has too closely modeled the experience of scientists moving forward in time on that of the historian moving backward. And that resulted in a number of difficulties: (i) the historian’s perspective generally encompasses a single step a number of changes which took place historically in smaller steps; (ii) it is generally a group of scientists that moves forward in time and an individual (the historian) who moves back, and the same descriptive arsenal cannot uncritically be applied to both. An individual can, for example, experience a gestalt switch, but it would be a category-mistake, in Kuhn’s view, to attribute such an experience to a group. It is mistakes of both these sorts which have made it harder than it should have been to describe the procedures available to the protagonist-scientists at times of theory-choice. For this reason, Kuhn now proposes to focus on the historian’s experience, as he presumably has already in the three examples mentioned above, and therefore it is explicitly in terms of the historian’s experience that he reintroduces the problem of incommensurability.

After recalling the etymological origin of the term from ancient Greek mathematics --where it specified the relationship between two quantities which had no common measure, no unit which each contained some integral number
of times, Kuhn says that the term 'incommensurability' is applied metaphorically to the relation between successive scientific theories; and it meant (using his present terminology): no common lexicon, no set of terms with which all components of both theories could be fully and precisely stated (see Kuhn, 1962; 1982; and also Part I, chapter 2). Today, Kuhn thinks that 'untranslatable' (Quine's term) is better word than 'incommensurable' for what he had in mind. Thus instead of saying that Aristotle's physics and Newton's physics are (here-and-there, i.e., locally) incommensurable, one should say that some Aristotelian beliefs are not translatable using a Newtonian lexicon or a more recent lexicon from physics. Accordingly, Kuhn points out, it is truth-preserving translation that cannot be done.

In Kuhn's view, the difficulties of translating science are far more like those of translating literature than has been generally supposed. They arise not only when translating from one language to another but also when translating between earlier and later versions of the same language. The three examples discussed earlier exemplify presumably these difficulties. In flagrant disagreement with the (realists') assumptions made by Kitcher (1978) and Newton-Smith (1981) for example, Kuhn argues in substance that in the absence of extended ethnographic interpretation --interpretation which transcends translation by summoning
unfamiliar meanings of some of the terms they contain—
each of his exemplary texts was systematically misleading.
Occasional passages, clearly of central significance for
their authors, patent ly failed to catch the meaning of the
passages they replaced. Some sentences in those passages
—sentences which must have been either true or false in the
original—today read so strangely in translation that it is
problematic whether what they appear to say can support
truth values at all. It is sentences or statements of this
sort that he has in mind when substituting untranslatability
for ‘incommensurability’, statements for which no
available translation techniques permit the preservation of
truth value. When he speaks of statements as translatable
or untranslatable, it is truth-preserving translation that
he has in mind (Kuhn, 1982/1983).

But let us ask: why can’t one assume that anything
which can be said in one language can be said in any other?
—especially if the lexicon of the language into which
translation occurs is suitably enriched. Kuhn’s response is
that if this thesis were correct, then anything said in one
language would carry its truth value with it when translated
to another. Otherwise a statement could be true in one
language, false when translated into another, and this would
lead to a type of relativism that he finds quite rightly
unacceptable. But he insists that another sort of
relativism is unavoidable. For example, in those cases where a statement which is candidate for truth or falsity in one language may, in another, be impossible to state as a candidate for truth value at all. In Kuhn's judgement, something of that sort is involved in the examples that he presented. [That is why he asked for example whether Aristotle was simply mistaken when he proclaimed the impossibility of void (1987)].

For Kuhn, though many of the statements that can be made with the lexicon of one language can be made also with the lexicon of another or with that of the same language at a later time, other statements cannot be carried over, even with the aid of an enriched lexicon. The content of those statements can be communicated nevertheless, but what is required is not translation but language learning (1982). That is what he had to engage in before he could understand the texts he discussed. Though he was thus able to communicate many of the beliefs of the scientists that he considered (i.e., Aristotle, Newton, Volta, Faraday, Planck), he insists that none of the terms that these scientists used (i.e., motion, matter, void, battery, electrical resistance, oscillator, energy element) applied to natural phenomena in the same way as their later replacements, and the lexicon in which those later terms occur cannot be used to provide words or phrases that can
substitute for them, salva veritates, i.e., in truth-preserving way.

In order to begin redeeming some of the promisory notes that I made so far on Kuhn's behalf, it is necessary to examine the model of lexicon that he develops. He considers a lexicon as it is embodied by individual members of a linguistic community and also as it is embodied in the language community as a whole. He cautions that though the model that he has in mind is simplistic and schematic, it illustrates nevertheless characteristics which any more articulated version should display. More specifically, he claims, it is particularly helpful concerning the difficulties of textual interpretations discussed earlier; and it constitutes also a sketch of the theory of meaning, in which the sense (or meaning) of a term is related to the way in which its referent is determined, without presumably falling prey to the objections that have typically been raised against early contextualist/verification theories of meaning.

Kuhn is concerned mainly with that part of the lexicon that contains putatively referring terms, each linked to the names of features that are of use in picking out its referents. In his view, the lexicon embodies the taxonomy of the language community whose members use it. Such a lexicon names the kinds of things, behaviors, and situations
which occur in their natural and social world, and it also names the more salient or central characteristics of those kinds, the characteristics by which they are known. The knowledge it embodies is thus about both language and the world, about the names of things and properties, on the one hand, and about those same things and properties, on the other. Those properties need not specify necessary and/or sufficient conditions for class membership, they simply provide useful guidance to classification. Better yet, they provide what Kuhn calls, a "feature space" within which elements, objects, bodies, entities, cluster like-to-like and contrast with others --close to them, or which could easily be confused with them. Thus, considering these properties collectively, one can say that any X is more like some other X's than it is like any Y.

There is a sense in which Kuhn could be taken to be advancing here a cluster theory of meaning. Because that "danger" exists, he gives three reasons for why such a theory differs from the one that he wishes to advocate. But in each case, one can show that it is in fact very much like the cluster theory that I will defend. First, he characterizes a cluster theory of meaning as a theory according to which an object belongs in a given category if and only if it manifests "enough" of that category's defining features. As we can see, his characterization is
exactly like the one given for the traditional versions of cluster theory (Part I, chap. 3/4). And, he points out that:

(i) In his theory, unlike in the latter, the properties do not presumably attach uniquely to individual categories; instead they serve to provide a space in which to determine the membership of whole set of inter-related categories which, for that reason, have to be acquired together. These properties provide information about properties that the members of a given category tend to share, but a more significant role is often played by properties with respect to which members of different categories differ. These differentiating properties, Kuhn claims, open "empty space" between the regions occupied by the members of the various categories. (see on this score Grandy's discussion of role and import of "semantic fields" and "contrast sets", 1986; 1990; see also Kuhn, 1982, p.14).

But, as I have suggested throughout, and as I will again stress in Part IV, a cluster theory (as I envisage it) need not assume or hold that properties attach uniquely to individual categories. In fact, I will argue it is possible to construe a cluster theory of meaning which satisfies what I have called the principle of correlation and "local holism." In that theory, clusters of (correlated) properties would serve to provide a "cluster space" within which to determine the membership of a Cluster of correlated
categories, which have to be learned together. Thus, the role and importance of "cluster semantic fields" and "contrast clusters" would be acknowledged.

(ii) More importantly, because of the open space mentioned above, Kuhn says, there is no need to specify the number of properties an object must possess or manifest in order to belong to a category. That is presumably the significance of the fact that we are dealing with terms marked in the lexicon as the names of natural kinds. Since, as Kuhn pointed out, that knowledge of nature is embodied in the lexicon, when that knowledge is threatened, or put into question, the lexicon itself, and not just particular beliefs which can be stated with its aid, is in jeopardy. This may even lead to "lexical re-design" or "taxonomic reshuffling." And this may involve for example:

- Different groupings (or cluster-configurations) of the objects or entities under consideration, with different similarity-difference relations holding between them.
- Inclusion or exclusion of one or more objects or entities.
- Addition or deletion of one or more new properties (e.g., more 'theoretical,' and less 'observational' or vice versa).
- Variations in the salience (or centrality) of the same or different (clusters of) properties.
- These and other kinds of changes or variations may be also associated with the introduction of a new instrument (e.g.,
telescope, microscope, balance, etc)—a point which has often been overlooked, with the new reading of an cluster-concept experimental setting, or with the formulation of a new theory, etc.

Let us note again that in the cluster theory that I envisage, one need not either specify the number of properties an object must possess or manifest in order to belong to a category. In part IV, I shall have more to say about these and other notions pertaining to the cluster theory that I have in mind. In the meantime, it shall suffice to note that much of what Kuhn claims for his theory can be expressed in terms of the cluster theory envisaged.

(iii) Finally, for Kuhn, what characterizes a lexicon is not the properties (or clusters thereof) it deploys, but the groupings (or rather clusters of correlated entities or objects) which result from the use of those properties (or clusters thereof), whatever they may be. And so what changes during scientific change (e.g., transition from ancient Greek to modern astronomy) is primarily similarity-difference relations posited by different lexicons about a set (or cluster) of objects (e.g., in the heavens).7 Think here of the relations holding between members of a cluster, and which serve to characterize the structure of the cluster in question; what are they, if not similarity-difference relations? (see Gasking, 1960; Part I, chapter 4 [A]). Kuhn insists rightly, the same groupings (or clusters) may be
achieved in many different "feature spaces." Even if, in principle, no two members of a language community need use any of the same properties (or clusters thereof) in order to classify the objects in their environment into the same clusters of natural kinds, in practice, many of the properties (or clusters thereof) that they use are most likely the same. Kuhn's answer in this respect would be that what is at issue is not whether they can identify the same properties (or clusters thereof) but whether they use the same ones in picking out the referents of a given term. This is not to say however that different individuals can use just any properties (or clusters thereof). According to Kuhn, two characteristics must be shared by members of a speech community if they are to divide the phenomenal world into the same natural kinds, identify the same objects and situations as members of those kinds, and employ these identifications in their interactions with the world and with each other. (a) The same terms must be tagged with the same natural kind label in their individual embodiment of the lexicon. (b) Whatever properties (or clusters thereof) their individual lexicon embodies, each of the "feature spaces" that results must yield the same hierarchical relations among kind terms and the same similarity-difference relations among the referents of terms at the same hierarchical level.
Thus, for Kuhn, the meaning of a term is associated, not with any particular set (or cluster) of properties, but with what Kuhn calls the "lexicon's structure," i.e., the hierarchical and similarity-difference relations that it embodies. And consequently, what separates two different languages or theories (e.g., those of ancient and modern astronomy) is not so much that they used different (clusters of) properties to pick out the referents of terms, but that those terms occurred in lexicons with relevantly different structures. Kuhn finally argues that it is structural difference in lexicons which might and often does impede truth-preserving translation.

In the end, Kuhn stresses that it is only the structure of the lexicon, not the feature space in which each community member embeds it, that need be shared. If and when such a shared structure is in place, each individual can then learn things that another individual knows, and they can both go on to learn new things about the world. But, for Kuhn, what one is committed to by a lexicon is not therefore a world but a set of possible worlds, worlds which share natural kinds and thus share an ontology. Discovering the actual world among the members of that set is what the members of scientific communities undertake to do, and what results from their efforts is the enterprise he once called normal science (1989).

The set of possible worlds open to members of a given
scientific community is presumably constrained and limited by the shared lexical structure on which communication between the community's members depends. But scientific development has sometimes had to push or break those limits, to "force" restructuration of some part of the lexicon and give somehow access to previously inaccessible worlds. This has happened a number of times in the history of science. In Kuhn's analysis, it is the occurrence of such episodes in the history of science which gives rise to what he calls "textual anomalies," and which he discussed earlier as incommensurabilities, cases of untranslatability. It is such anomalies, Kuhn claims, that truth-preserving translation cannot remove. But he adds, where translation fails, there is another recourse, and that is, language learning. This recourse does not license however, let alone justify the use of such terms as 'true' or 'false'. Interestingly enough from Kuhn's point of view, (and contra Kitcher's and Newton-Smith's), the role of such terms is intra-linguistic or intra-theoretical, i.e., confined to the evaluation of choices between the worlds to which the structure of the community's lexicon gives access.

By way of partial summary, let us recall that Kuhn argued that a historian must undertake a quasi-ethnographic interpretation in order to understand a body of past beliefs, to make it plausible and coherent; three sort of
examples are examined (1987) to illustrate this point. He then stressed that the need for such an interpretation arises from a disparity between the taxonomy current in the historian’s linguistic community and that of the other community for which the texts he studies were intended. As we have seen, according to Kuhn’s model of the lexicon, the lexicon of natural kind terms provides taxonomic categories and ways of applying and using them. Taxonomic disparities between past and present sometimes emerge however; these disparities preclude truth-preserving translation of some central past beliefs into modern terms. But as Kuhn pointed out repeatedly, the primary and indispensable role of judgements of truth and falsity was played within history, not across it. Terms like ‘true’ and ‘false’ need function only in the evaluation of the day-to-day choices made within a community that has an ontology of kinds and a corresponding lexicon in place. He always insisted that these barriers to the evaluation of truth claims do not necessarily constitute by the same token barriers to understanding. Since understanding is the goal of the historian, what then must he do? According to Kuhn, he needs to acquire the lexicon of the community under study; that is, he must assimilate its taxonomy of natural kinds, and "otherwise situate himself imaginatively in its world." But Kuhn stresses that he need not and cannot properly
employ his own lexicon and the knowledge of his own world in a piecemeal evaluation of the truth-claims of the older (or other) community. Though there are judgements he can make about the past, they cannot be properly formulated in terms of 'truth' and 'falsity'.

Finally, Kuhn confronts some of the problems which may be encountered by the theory of meaning he sketched out. In anticipation, he argues essentially that the problems often raised against his position are either not pertinent or else not appropriate causes for alarm. This concurs with the underlying line of reasoning that I have taken throughout with respect to Kuhn's work. And so, accordingly, his answers could also be drawn on account of the cluster theory that I will advocate.

As Kuhn sees it, his position may be faced by four problems. Let us consider (1) the problem of bridgeheads. Kuhn has argued that past beliefs are recaptured by language learning, i.e., by acquiring the lexicon of kind terms in which those beliefs were stated. Now there may be cases where two lexicons are so disparate, i.e., "incommensurable," that there is no way of moving from one to the other. But Kuhn dismisses the possible threat that such cases can pose by pointing out that where two community-based languages are involved, it is very unlikely that they will happen very frequently. Speakers of one human language seem able always
to find a bridgehead from which to enter another, and some such a bridgehead is essential to the acquisition of a second lexicon. The required bridgehead, in Kuhn's view, need not be particularly broad or solid. The only condition placed on it, if it is to serve its function, is that some of the taxonomic categories provided by one lexicon must overlap substantially in their membership with categories in the other. In other words, only overlap is required, not identity of membership.10

In order for an individual to acquire a second lexicon, s/he must be able to form and test hypotheses about the particular entity or event to which a user of that lexicon is referring when employing a particular term.11 Some of the entities or events grouped together by the new lexicon must be grouped together also by the old or other one. According to Kuhn, these are conditions only on the referents of lexical items; that they are satisfied tells nothing at all about what those items mean. The bridgehead need not supply any constraints at all with respect to meanings. For Kuhn, it takes another, a second process to establish meanings: The individual must somehow find properties (or clusters thereof) shared by the various occasions when a given term was used as well as properties (or clusters thereof) which differentiate those occasions from others when the individual anticipates the same term
but does not encounter it." (ref: semantic fields and contrast sets). But as Kuhn also stated, no two individuals need select the same properties, but each must select properties (or clusters thereof) which yield the same taxonomy, the same lexical structure, the same similarity-difference relations among the referents of natural kind terms. Otherwise they will pick out different referents for the same terms, and communications involving those terms will break down. Only differences in lexical structure can limit the possibility of truth-preserving translation. But this does not preclude the possibility of understanding a different lexicon, with a different structure.

Since to learn new features is to learn new ways of discriminating, Kuhn assumes that the shared biological heritage of homo sapiens makes it natural to suppose that a discrimination used by any normally equipped human being can be learned by any other. And it is presumably this universal learnability of properties (or clusters thereof) which guarantees the existence of bridgeheads from which differently structured lexicons can be acquired. [Doesn't this lend support to the principle of humanity?] 12 It is one thing however to consider the existence in principle of bridgeheads and quite another to consider what might obtain in practice. In historical development for example, major mismatches of lexical structure may obtain. But even these,
they are to be expected, according to Kuhn, only when comparing lexicons from very different periods or cultures, e.g., antiquity and the 17th century, and even these mismatches are not total. In any case, Kuhn points out, they are of concern only to the historian who is looking backwards across them to the starting-point of his narrative. But as his narrative moves forward, within the developmental process itself, structural changes take place in smaller, isolatable stages, and are often restricted to "local regions" of the lexicon (e.g., force/mass/weight; battery/resistance/electrical current; oscillator/energy element). This is to say that outside these localized regions, structures are homologous, and truth-preserving translations possible. In a way reminiscent of the assumption made by Shapere/Nersessian in their contention about "chain-of-reasoning" connections (CORC's), Kuhn believes that bridgeheads in actual historical transitions are strong.

Concerning (2) the problem of relativism, Kuhn remarks that we still have to deal with those regions of the lexicon in which structure differs, and which contain the fundamental terms of a science --terms which like those just mentioned, are required to state its constitutive generalizations and laws. And as Kuhn points out, these generalizations are not translatable, like other statements containing those terms. For modern knowledge cannot be used
to judge them true or false. Truth values applicable to
them can be supplied only from within the lexicon used to
state them. Indeed, Kuhn insists, many statements
constitutive of a science can be assigned truth values only
from within the community of practitioners. This makes his
position seem relativistic. But Kuhn begins by dispelling
such an objection by first reformulating the question: the
question, he claims, should not be: is his position
relativistic or not, but rather "is anything worth
preserving lost?" by adopting such a position? According to
Kuhn, the answer is 'no', for it is not truth value that he
is relativizing, but only what can be said, i.e., effability
(1989). Kuhn's point is that if the same statement can be
made with the lexicon of different communities, it must have
the same truth value in all. But in point of fact, some
statements which are clear candidates for truth value in one
community are simply unsayable in another. The entities or
events they describe do not occur in any of the possible
worlds to which that community's lexicon gives access. It
is such statements that Kuhn claims are impossible to
translate with the specificity required by the truth-value
game. For Kuhn, the evaluation of a lexicon with which to
describe phenomena and to build theories about them is a
very different undertaking from the assignment of truth
values for the individual statements that the lexicon
permits one to construct. For Kuhn, evaluating a lexicon is a different matter, for a lexicon cannot properly be labelled true or false. Its structure, the taxonomy it provides, is a matter of social and linguistic fact. Nor can a lexicon properly be described as "confused," though its use may on occasions result in confusion. Instead, one lexicon is a better or worse instrument than another for achieving specifiable social goals, and the choice between lexicons --and even the direction of lexical evolution-- necessarily depends upon those goals. What Kuhn may be suggesting here is that the pragmatists were generally right in this respect. Lexicons are instruments to be judged by their comparative effectiveness in promoting the ends for which they are put to use. The "choice" between them is interest-relative. If Kuhn's position is instrumental and relativistic with respect to lexicons, it does not follow that his position is relativistic with respect to truth. This point is apparently crucial to Kuhn's position. Even if Kuhn's position still has some difficulties, relativism as it is commonly understood is presumably not one of them.

Turning to (3) the problem of realism, Kuhn raises the same old skeptical questions, and undermines thus the bases for scientific realism. In his view, truth claims and their evaluation remain in place where they are needed, and that is, within a community's day-to-day life. However,
the evaluation of truth claims is often impossible between communities with different lexical structures. But then, as we have seen, Kuhn claims, on the basis of our common humanity and universal dispositions, that there are almost always sufficient bridgeheads available to permit judgements about, and thus some understanding of, the way of practising science that the other community’s lexicon permits. If that way is found to be better in some respects (e.g., if it offers solutions to yet unsolved problems; if it has a greater problem-solving capacity or greater explanatory power), then one may choose to adopt it, and thus, enter a new community, acquire its lexical structure, and learn its language. Metaphorically at least, that is the way much of scientific progress occurs, for Kuhn. And the reality of progress is not for him in doubt. (see also SSR, Postscript, 1970, p.206). But that progress is instrumental. Though real, it is not progress towards reality. Science does provide us with a more and more powerful taxonomy for dealing with the world, but it does not do so by discovering a lexicon-independent truth.

We have already seen that, in view of the difficulties that it raises, Kuhn rejects the notion of inter-theoretic truth. But, in previous writings (1962;1970;1977), Kuhn was also suspicious of the idea of a theory-independent truth. Now since theories are possible only by virtue of their
dependence on a lexicon, and are thus in an intermediate position, Kuhn expresses the same idea, but this time with the notion of lexicon: there is no lexicon-independent truth. By the same token, he was also, and still is, suspicious of the idea that scientific progress consists in getting "closer and closer to the truth" or else "zeroing in," "cutting nature closer to its joints." In this respect, i.e., the ontology embodied in the lexicon, Kuhn points out, the history of science displays nothing like "zeroing in." While problem-solutions advance steadily in number and precision, the ontologies from which those solutions derive vary widely in numerous directions. To date no one has shown how to display anything like an asymptote towards which science, throughout its history, has been moving closer. At any time, there are best scientific guesses about the nature of the world's ultimate parts, but they do not stay in place as lexicons change.

Similar and related considerations in SSR (1962) had led Kuhn to talk of the world's changing during revolutionary episodes in a science, or else to say that, after a revolution, "scientists live and work in a different world" (see 1962, pp.111ff;150; see also Part I, chapter 2). These expressions were, if you recall, perceived to be rather 'puzzling'. Now Kuhn realizes why such expressions were misleading, even though appropriate in some sense, and
he suggests a new way for dealing with them. What is misleading for Kuhn about saying "the world changes" is the implied idea that the community stood still while the world changed about it. In fact, Kuhn now stresses, both the world and the community changed together with the change in the lexicon through which they interacted. He helps us better understand what he is saying by considering the example of 'phlogiston' in 18th century chemistry (1982; 1987). Kuhn's view can be schematically represented as follows. [These schematic representations might make easier to comprehend the relations that Kuhn claims hold between a community, its lexicon, and the world]:

```
(a) Community ----> Lexicon ----> World
    
(b) Lexicon -------------> World
    /    \                  /    \
   <Structure>      <Groupings>
   <Inter-related Terms> <Clusters of Entities>
   <Clusters of Properties> <Objects, Events, etc>
```
Finally turning to (4) the problem of connections between past and present, Kuhn begins by breaking it down into two questions: (a) Given the role of lexical structure in constituting a world, how can a lexicon be changed? (b) Given the problems of translatability which result from lexical change, what connections with the past are available to the present; how can the past be recaptured, and made part of the present?

Given Kuhn's view, one can assume his response to (a) to be as follows. The community's knowledge of the world is built into the structure of its lexicon; new data and experiences sometimes "strain" that built-in knowledge in ways that can only be relieved by lexical change. Such strains can come about in various ways. Kuhn often cites, for example, the strain placed on Aristotle's concept of 'motion' by Galileo's experiments with an inclined planes, the strain placed on Newton's concepts of 'space' and 'time' by Einstein's famous thought experiment about the moving train struck by lightning at both ends. Though these examples are not typical, and though there are other routes
to lexical change, they have presumably "the special virtue of displaying clearly the lexicon's involvement in revolutionary change of theory. By doing so they locate the strain that precedes such change where it belongs — on the lexicon" (1987).

As for question (b), it comes down, according to Kuhn, to figuring out how historical narratives bridge the gaps or ruptures that remain after lexical change? In his view, this question has two answers, both inescapable and mutually deeply inconsistent, and which reflect two different conceptions of history (1986): (i) The first answer presupposes the sort of history that Kuhn has been advocating throughout his career. In this case, a historical narrative begins with a quasi ethnographic or hermeneutic reconstruction of the relevant aspects of some period in the past. Such a narrative sets up the context, introduces the main individuals involved, their lexicon, their world, and finally describes the strains on the lexicon as they arise and accumulate, thus making way for a new lexicon. (ii) The second answer presupposes the sort of history which is known as 'Whiggish history'. In this case, the historian does not attempt to acquire and deploy the older lexicons used by the community under study; rather, he seeks to tell the story of the past in a translation which does not attempt to preserve truth values,
but rather a constant truth, that emerging on the historian's horizon. Even though throughout his career, Kuhn has consistently denounced and ridiculed such kind of history, he now seems to think that 'Whiggish history' has an indispensable human function which cannot be fulfilled by the first sort of history.

The two narrative modes described above are clearly never found in pure form; to a greater or lesser extent, they always interpenetrate. But the cleavage between them is nonetheless real. In a more muted form, one might say that it separates historians whose subject is their native land from those who study other nations or other cultures. As Kuhn seems to suggest, it is the cleavage between those who need history to look back and those who need it to look ahead, and it will not be eliminated. This is I think an interesting new development in Kuhn's view, which has wide ranging implications, as we have seen. We should keep it in mind when turning to the new and revised cluster theory that I will advocate next in Part IV. In anticipation, I will assume that the conception of history underlying the approach advocated by this cluster theory is of the former kind, i.e., needed for looking back into the past of the scientific process.18 As for looking ahead into the future of science, we do not need a conception of history which assumes the truth of the current standpoint, and which
evaluates all of the past of science in relation to it. Instead, in our attempt to articulate the proper "historicist turn," we shall consider what we have learned from the past of science about science, and, if need be, use our "philosophical or scientific imagination" consistently with what we have learned.
Part IV: A NEW AND REVISED CLUSTER THEORY

In this part, I would like to articulate finally the new and revised cluster theory of the process of meaning-making in science that I have announced so far in bits and pieces. Though it is far from being complete and fully developed, it shall be sufficiently fleshed out to permit appreciation of its distinctiveness and evaluation of its adequacy. I will attempt to make the best possible case for that theory, while remaining open to the possibility that it may prove after further analysis to be inadequate, at least in some respects. But first, it is only appropriate to reformulate the general argument underlying this whole project, and situate the approach of the theory that I advocate with respect to all other approaches discussed in previous parts and chapters. This should enable me to characterize more explicitly than I have done thus far some of the main theoretical and methodological assumptions and principles of the new cluster theory. I will then further bring out the distinctive features of this theory by showing how it meets the objections raised against other accounts, and satisfies (most, if not all) the adequacy requirements formulated by these accounts. I will also show how it could lend itself to further extensions and refinements in light of recent advances in the cognitive sciences (e.g., cognitive psychology, AI) and in sociology and history of science. To begin testing the viability and
plausibility of the new cluster theory, I will turn (chapter 2) to a particular case-study, and apply it to the "revolutionary change" which led from Stahl's and Priestley's phlogiston theory to Lavoisier's oxygen theory in 18th century chemistry. Finally (in chapter 3), I will consider a few residual problems which may still be encountered by the new cluster theory, and which should be of concern to anyone interested in pursuing this line of research.

Chapter 1: Toward a Theory of the Process of Meaning-Making in Science

A. Motivation/Context

As I stated in the introduction to this project, my main goal is to make the best possible case for a new and revised cluster theory --as a potentially adequate theory of the process of meaning-making in science. But how best to make such a case? --if not by showing, as I think I have (though not in this order) the following. (i) There are no a priori grounds for not considering the theory here in question as a viable and plausible alternative. (ii) All or most major alternatives considered fail for some reasons or others. (iii) The most significant contributions of some of the recent and promising accounts can be incorporated in the theory in question.

Thus, in Part I, I have shown that (a) traditional contextual theories of meaning (Carnap [1956], Kuhn, [1962])
lead to the problem of extreme and radical incommensurability, because of their underlying "essentialism of meaning" and inadequate concept-representation. 1 (b) Traditional versions of cluster theory (Gasking [1960], Putnam [1962], Achinstein [1968]) fare better relatively speaking. They do so mainly because of the "cluster" device introduced presumably in order to take care of the flagrant essentialism of meaning underlying traditional contextual theories, and to provide a more adequate concept-representation— one which is not as rigid and insensitive to the temporal and developmental dimension of concept formation in science as the traditional concept-representation in terms of sets of necessary and/or sufficient conditions. But, as I pointed out, they fail with respect to the problem of incommensurability, and more generally as accounts of the process of meaning-making in science, for two sets of reasons: not just because of the traditional objections raised against such theories, but because they still assume some grade of essentialism (see discussion in Part I, chap. 3 of "grades of essentialism") and characterize the cluster-structure in terms of problematic or incomplete principles (e.g., omnifocality, over-determination, and semantical relevance). 2 (c) The assumptions and principles of these cluster theories are not necessary or inherent to a cluster theory.2 A new and
revised cluster theory can be formulated with a new set of assumptions and principles (both theoretical and methodological). (d) The objections commonly raised against a cluster approach, mostly by proponents of the reference approach, who argue subsequently that their approach is the only alternative left, can well be met by the new cluster theory. 3

In Part II, (a) I have shown why a reference approach in general (Scheffler [1967], Putnam [1973, 1975]) does not solve or dissolve the problem of incommensurability, and as a result, cannot provide an adequate account of the process of meaning-making science mainly because of its underlying "essentialism of reference" and its inability to account for reference changes. (b) I have shown why even an improved version of Putnam's CTR (i.e., Kitcher [1978, 1982]) will not do either --despite making some interesting points (e.g., Kitcher's notion of "reference potential"). (c) The reference approach is not the only alternative; furthermore, it also fails to do satisfactorily (i.e., without engendering serious new problems), what it claims to be able to do better than any other approach, including the cluster approach.

Finally, in Part III, I have attempted to show that the more recent and promising proposals of Shapere [1984, 1989], Nersessian [1984, 1987, 1989], and Kuhn [1988] have made some
significant contributions which must be taken into account by any adequate theory of the process of meaning-making in science. Furthermore, I have argued that the cluster theory that I envisage can well incorporate these contributions, or else, that these proposals are in fact best construed as putting forth or supporting such a cluster theory --since their proponents are arguably "closet cluster theorists," who deny that they are, mainly because they only have in mind traditional versions of cluster theory.

But as we shall see, there can be a new and revised cluster theory, different in many respects from the traditional versions, which constitutes a viable and plausible alternative, and which provides an adequate account of the process of meaning-making in science. In general terms, it would be fair to consider that such a theory constitutes in some sense an extrapolation and improvement upon Achinstein's version --but without the principle of semantical relevance. In addition, it manages to incorporate the main insights and contributions of the recent and promising proposals made by Shapere, Nersessian, and Kuhn. In passing, one should note that it is thus capable of accounting for whatever can be accounted by the improved version of Putnam's reference approach (i.e., Kitcher's) without encountering any of the major problems plaguing the latter. But in some other sense, it is somehow
so different (in its theoretical and methodological assumptions) from traditional versions of meaning/descriptivist/cluster theories that it might seem pertinent to ask if it is a cluster theory only by name. I shall postpone my answer to this question until I have had a chance to expose more fully the new cluster theory (see forthcoming chapter 3, on residual problems).

The following schematization might help situating (and perhaps better understanding the rationale behind) the new cluster theory in question:
THEORIES OF THE PROCESS OF MEANING-MAKING IN SCIENCE

<table>
<thead>
<tr>
<th>THEORY-DEPENDENT / TRANS-THEORETICAL ACCOUNTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I) Meaning Approaches</td>
</tr>
<tr>
<td>(a) Contextual Theories</td>
</tr>
<tr>
<td>Carnap [1956]</td>
</tr>
<tr>
<td>Kuhn [1962, 1970]</td>
</tr>
<tr>
<td>Problems of Incommensurability</td>
</tr>
<tr>
<td>(II) Reference Approaches</td>
</tr>
<tr>
<td>(a) Context-insensitive Theories</td>
</tr>
<tr>
<td>Scheffler [1967]</td>
</tr>
<tr>
<td>Putnam [1973, 1975]</td>
</tr>
<tr>
<td>(III) Cognitive/Historical Approaches</td>
</tr>
<tr>
<td>Shapere [1984]</td>
</tr>
<tr>
<td>Nersessian* [1984, 1987]</td>
</tr>
<tr>
<td>Kuhn [1988]</td>
</tr>
<tr>
<td>(b) Traditional Cluster Theories</td>
</tr>
<tr>
<td>Gasking [1960]</td>
</tr>
<tr>
<td>Putnam [1962]</td>
</tr>
<tr>
<td>Achinstein [1968]</td>
</tr>
<tr>
<td>(b) Context-sensitive Theories</td>
</tr>
<tr>
<td>*Field [1973]</td>
</tr>
<tr>
<td>Kitcher [1978, 1982]</td>
</tr>
<tr>
<td>*Newton-Smith [1981]</td>
</tr>
</tbody>
</table>

(IV) New Cluster Theory

(I) [(b)]: attempt to deal (unsuccessfully) with the problems of incommensurability to which (a) have led. (II): attempts to show why a meaning/cluster approach cannot succeed in avoiding/solving problems of incommensurability, and why focus on a reference approach is more promising. Under (II) (b), we could have also included * marked. (III): attempts to show why neither (I) nor (II) can succeed, and why a new approach must be tried. [*Some might object to Nersessian being classified here, and would argue that instead she should classified under (I) [(a) or (b)]. (IV): attempt to show that the new approach to explore is the one proposed by the new cluster theory —which is strictly speaking neither a meaning nor a reference approach, but a little bit of both, with a heavy injection of the adequacy requirements formulated by Shapere, Nersessian, and Kuhn.4

Thus, as is expected, the new cluster theory makes some
theoretical and methodological assumptions along the lines of Shapere, Nersessian and Kuhn, and supports by the same token some of the main tenets of the "New Philosophy of Science" (see Introduction). It takes the "historiocist turn" in its approach, in that it strives to incorporate as much as possible the "context of discovery." Science is considered to be an open-ended, on-going activity, whose character, standards, problems, and methods have changed and can still change significantly in its history (tenet #1): that is why the focus has been and still is, on the process of meaning-making in science. Science is not a monolithic enterprise (tenet #2): consequently, we should be weary of any sweeping generalizations which purport to apply to the "whole of science." Science has its roots in everyday circumstances, needs, methods, concepts, etc, of human beings (tenet #4): our theories of science should not postulate cognitive processes among scientists which are radically different in nature from those encountered and exemplified by ordinary cognitive agents; they should assume instead that there is a continuity between these cognitive processes, and that they can be mutually illuminating (see Giere, 1988; Nersessian, 1987; 1988; Tweney, 1981). It is through examination of actual science that we will learn about how to understand 'concept formation and development', 'theory', 'meaning', 'reference', 'explanation',
observation', 'progress', 'rationality', etc (tenet #5). In this respect, one ought to keep in mind Shapere's famous slogan: "We have learned how to learn" in studying science. The function of philosophy (of science) is not only descriptive -- at least initially, it is also critical and prescriptive (tenet #7) [see Chokr, 1988, pp.78-88]. It will also be assumed that the boundary between a 'scientific' and a philosophical question, especially when it comes to foundational problems is often blurry, and hard to justify (tenet #6) [see Quine for a similar point]. Finally, it will also be assumed that the philosophy of science is not self-sufficient (contrary to what has been assumed and claimed by the positivists, for example) -- insofar as the methods of philosophy need to be supplemented by concepts, principles, models, theories, etc drawn from other disciplines (e.g., cognitive sciences, history of science, sociology, linguistics, etc), even though it remains unclear or rather a matter of debate just what can be borrowed or transferred from these other disciplines and how it can be used [more on this point at a later point].

Concerning the problem of incommensurability, it seems to me that the mere contemplation of this problem has tempted so many able philosophers into a swamp of confusion. But as I have shown in Part I (chapters 1 and 2) and
throughout, incommensurability is not as threatening to the rationality and progress of science as it has been presumed. First, it is prudent to begin by distinguishing different sources and therefore different kinds of incommensurability, and furthermore, of various degrees. Second, even the most interesting and serious kind incommensurability (of meaning), [which presumably implies that they are no "theory-neutral" grounds for deciding what the proponents of two different scientific theories are referring to with their respective terms, and therefore, there are no rational grounds for comparing and choosing between them], we have seen that it is the consequence of an inadequate conception of meaning. In other words, incommensurability does not imply incommunicability, or incomparability. Such an implication might arise however if one adopts a holistic and essentialist theory of meaning. In such a theory, the meaning of all scientific terms is determined by the theory in which they are used, and the meaning of each is characterized in terms of necessary and/or sufficient conditions for its application, or else, in the traditional sense, in terms of a cluster of properties "enough" of which must be satisfied for any given term to be applicable. But as I will argue again in a moment, we need not adopt such a theory, which posits an 'infectious holism' and an unjustifiable 'essentialism of meaning'. Nor do we need to
adopt a reference approach (a la Putnam), which seeks to insure continuity of scientific discussion and overcome any problems of incommensurability of meaning by focusing on reference, and by positing (almost arbitrarily) a "rigidity of reference" and "essentialism of reference" (Part II, III).

Having said this, I don't mean to suggest that there is no real problem for philosophers of science to deal with concerning scientific change. Again, as I pointed out in Part I, we have learned from the history of science that different scientific theories talk in sometimes very different ways about different sorts of entities, and even what those theories are about may seem or be different. If two scientific theories differ in these and many other respects, as they certainly seem to do, on what basis and how do we compare them? The comparison of scientific theories is a real problem, contrary to what the positivists may have assumed, but there is no in principle argument against its possibility. Rather than concluding that we cannot, and thus threaten the rationality and continuity of scientific discussion, on the basis of some ill-begotten theory of meaning, which posits an "essentialism of meaning"; and rather than concluding that we can, and thus insure somewhat arbitrarily the rationality and continuity of scientific discussion, on the basis of some ill-begotten theory of reference which posits an unjustifiable
"essentialism of reference," we need to ask: how we can actually compare scientific theories, and how in fact we do make such comparisons? And subsequently, we need to explore seriously the possibility of providing an account which tells this story, while not assuming any form or grade of essentialism (whether of meaning or reference), and being thus in accord with the open-ended nature of the scientific process.

Assuming the evolution of science, i.e., a connection or traceable line of descent between scientific theories (cf: Kuhn, 1970; Shapere, 1984), what we want is a theory which is able to account for continuity, commensurability and comparability despite discontinuities or incommensurabilities --even of the most radical kind, as a result of revolutionary change. For even in such cases, there is arguably still continuity. [see forthcoming case-study of "chemical revolution"]. What we need is a theory which can account for the various and different kinds and degrees of changes which take place within the process of meaning-making in science. These changes might include minor meaning-changes --which account for most of scientific development (e.g., in the cluster of properties associated with a term and ascribed to an entity or type of entity), or radical changes (as when a term, its associated cluster of properties, and the entity or type of entity it referred to
in a previous theory are abandoned in a later theory). What we need is a theory which accounts for continuity and comparability in science — not by common ascriptions of (essential) properties (or clusters thereof) to the terms used by proponents of different scientific theories, nor by asserting (arbitrarily) commonality of reference despite theoretical change, but by tracing out, using Shapere’s CORC device, the reasons given by scientists for changing the cluster of properties associated with a term in a given theory, and/or for changing and even abandoning the referent(s) that this term picked out.

I propose that an appropriately reformulated cluster theory, like the one I envisage, constitutes such an account. I will show that such a theory is different in a number of respects from traditional versions of cluster theory, and that it satisfies various other adequacy requirements formulated by Shapere, Nersessian, Kuhn, and others. And furthermore, that when it is supplemented with a sophisticated semantical apparatus, like I intend to do, it enables us to give a fine-structured and context-sensitive reading of the history of science.

At this point, it seems then appropriate to articulate more explicitly and systematically the major tenets and principles, in short, the distinctive features of the new cluster theory by contrasting it with the traditional
versions of the same kind. I will distinguish for my present purposes between a general traditional version (Part I, chap.3) and specific versions (Gasking, Putnam, Achinstein).

General Traditional Version of Cluster Theory:

T1: The meaning of a term is a function of a monolithic and determinant theory or theoretical background in which the term is used.

T2: The meaning of a term is determined by a cluster of (relevant) properties, "enough" of which must be satisfied for the term to be applicable to a given entity or type of entity.

T3: The cluster of (relevant) properties constituting the meaning of the term determines what the term refers to.

T4: An individual is said to know how to use a term (and to refer to what it is supposed to refer to), iff s/he can provide and knows the cluster of (relevant) properties constituting the meaning of the term in question, and associated with it.

There are clearly remnants of 'essentialism of meaning' in this version of cluster theory, as indicated by T1 and T2. The nature of the relationship between "background" and the meaning of a term is all-too deterministic. The insistence (in T2) on a number of properties [enough, most, many] without specifying what would be satisfactory leaves open the question of how far wrong can the properties in the cluster or the cluster of properties be and still succeed in referring to what it is supposed to pick out. Furthermore, the nature of the relationship between the cluster of properties constituting the meaning of a term and its
referent is too-strict and all *too-determinant* [T3]. Finally, we should note as T4 indicates that the *locus of meaning* is the *individual* (rather than the *community*), and that *linguistic competence*, or simply use of a term, presupposes *epistemic access* on the part of the individual.

**Specific Traditional Versions of Cluster Theory:**

On the whole, it is fair to say that Gasking, Putnam, and Achinstein make claims similar to T1-T4, except that in addition, each puts forth a different principle to characterize the structure of a cluster of properties, or its relation to the things it refers to. Thus, for Gasking, the following holds:

**G1:** The 'meaning' of a term is always to be understood relative to a *determinant* (theoretical) background context.

**G2:** The 'meaning' of a term is *determined* by a cluster of *omnifocal* properties, *most* of which must be satisfied for the term to be applicable to a given entity.

**G3:** The cluster of *omnifocal* properties constituting the 'meaning' of a term *determines* what the term refers to.

**G4:** An individual is said to know how to use a term (and to refer to what it is supposed to refer to) *iff* s/he can provide and knows (i) *one* (or a few) of the properties in the cluster (of *omnifocal* properties) constituting the 'meaning' of the term in question; or (ii) (one of) the relation(s) holding between the properties in the cluster.

In proposition G4, Gasking seems to have the beginning of the correct answer to the question implied here: does an individual need to know, and thus have epistemic access to, the cluster of properties constituting the 'meaning' of a
term in order to be able to use this term successfully? If this were the case, then we would have to say that you and I do not know how to use very many terms (scientific or ordinary). This would be rather counter-intuitive, apart from being incorrect. The correct answer, as we shall see, consists in saying that an individual need not have the postulated epistemic access and explaining why s/he may still be said to know how to use a term, and many others. Gasking seems to lean toward the first part of the answer (see Part I, chapter 4 [A] for details). But when it comes to second part, he misguidedly offers 63 on the basis of principle of omnifocality, which serves him in characterizing the structure of the cluster of properties. If you recall, what Gasking means by such a principle is that the properties constituting the cluster are each a potential focus, of equal weight or value and all so strongly related that it suffices to know one (or a few) of the properties and (one of) the relation(s) holding between them to be able to retrieve the whole cluster. But, as I argued already in Part I, Gasking’s principle is unnecessarily too strong: the properties constituting a cluster cannot be on a par, of equal weight or value, and equally a focus; all we need at this point is a principle of correlation, whereby the nature of the relation(s) holding between the properties in the cluster is specified to the
extent that it is possible, and whereby the properties in
the cluster which vary over time are of various and unequal
weight or value, some more salient, typical or central than
others. Rather giving in and accepting some form of limited
epistemic access, like Gasking, we need to stress that under
an appropriate construal of cluster theory, no epistemic
access is required— even though some form and degree of
linguistic competence (actual or 'parasitic') might be in
order [cf: Putnam's principle of the division of linguistic
and cognitive labor (1975),6 and the corollary principle of
derference to relevant experts and authorities].

As for Putnam (1962), he would make the following
claims:

P1: The 'meaning' of a term—as reference-fixer, is always
to be understood as relative to a monolithic and determinant
background context.

P2: The 'meaning' of a term is determined by a cluster of
(relevant) properties (laws), most/many/a good number of
which must be satisfied for the term to be applicable to a
given entity.

P3: The cluster of properties constituting the 'meaning' of
a term over-determines what the term refers to.

P4: An individual is said to know how to use a term (and
refer to what it is supposed to refer to) iff s/he can
provide and knows the cluster of properties constituting the
'meaning' of the term in question.

By virtue of what I have called the principle of
overdetermination, Putnam states P3: the cluster of
properties over-determines the referent. Thus, as a result
of theory change, one or two (may be more) properties (or laws, rather) may be added or deleted from the cluster of properties which in a previous theory constituted the 'meaning' of a term, and still refer to the same entity. But, as I pointed out in Part I chapter 4 [B], the principle of under-determination holds as well in the following manner: the referent of a given term under-determines the cluster of properties associated with that term; the same referent can still be picked out by (slightly) different clusters of properties.

Finally, as for Achinstein, he would hold:

A1: The 'meaning' (as use) of a term is characterized on purely linguistic grounds as a function of a monolithic and determinant 'background context,' relative to a given community.

A2: The 'meaning' of a term is determined by a cluster of semantically relevant properties, most/many of which must be satisfied for the term to be applicable to a given entity. Semantically relevant properties are those properties which are relevant in and of themselves; in other words, which are irreducible (Achinstein, 1968, pp.8-9;25; Part I, Chap. 4).

A3: The cluster of semantically relevant properties associated (on linguistic grounds alone) with a term over-determines what the term refers to. Different cluster-configurations may still pick out the same referent(s).

A4: An individual is said to know how to use a term (and refer to what it is supposed to refer to) iff s/he can provide and knows the (semantically relevant) cluster of semantically relevant properties constituting the meaning of the term in question, and associated with it on linguistic grounds alone.

Apart from all the other problems and difficulties plaguing Achinstein's account, proposition A4 reveals
clearly what I have suggested in Part I, Chapter 4[C], i.e.,
the vicious circularity which arises from his use of a
semantical notion to explain a semantical phenomenon. In
the end, Achinstein does not explain anything with the
principle of semantical relevance, because he says, in
effect, that 'to know how to use a term' is 'to know how to
use a term' (see Part I, chap.4 [C]).7 I will therefore
argue that a better, non-circular, principle in this respect
is what I shall call the principle of assertibility
(relative to a given "multivariate background context,"
which itself is relative to a given community and its
members). [More on this next].

The various cluster theories of meaning which can be
extrapolated from these discussions can be schematically
represented as follows:

Background
Context (Theoretical)

\[
\begin{array}{c}
\text{Term: Meaning or Referent(s)} \\
\text{'Meaning'} \\
\text{--as reference-fixer} \\
\text{--as use} \\
[\text{enough/many/most (omnifocal/relevant/)}] \\
[\text{semantically relevant) properties}] \\
\text{[associated with term]} \\
\text{[on empirical grounds (Gasking)]} \\
\text{[on linguistic grounds (Putnam/Achinstein)]}
\end{array}
\]

[more to follow]
show in what senses it constitutes a radical improvement. I will argue that, in contrast with previous theories, it is based on more plausible assumptions and principles, and is thus better equipped to meet the objections raised against traditional theories of meaning.

The cluster of properties associated with a term does not determine or constitute its meaning, even though such a term does have meaning ["meaning as use" --'meaning']. Some of the properties in the cluster may be part of the meaning of the term, while others may not be. Some of the properties used to introduce a term may turn out not to be significant at all. The primary function of the cluster of properties associated with a term is to characterize theoretically and to secure reference, not to determine meaning, i.e., the semantic characterization of the things referred to by the term. This is by the way one of the features of the new cluster theory which may prompt one to ask whether this is not a cluster theory only by name. Under this theory, clusters of properties are associated with terms, as are references. But the relationship between term, associated cluster of properties, and reference is not that the associated cluster of properties determines the reference. Nor is it that the cluster of properties is associated with a term on linguistic grounds alone. In fact, it is important to stress that the three somehow
interact more flexibly. An associated cluster of properties can be generated from the reference. The discovery of entities similar to those in the reference, but lacking the associated cluster of properties can weaken the association of a property with a term. Similarly, the discovery of things different from those in the reference but having the associated cluster of properties can weaken the association of a property with a term.

One might object that at any point in time there are degrees of association for what we might call "unreduced properties," and that this constitutes in effect the meaning or semantical use of the term at that point in time. Let us suppose that there are such degrees of fixed (or temporarily stable) association. Are those associations mere linguistic facts, or else to be explained on linguistic grounds alone? Contrast the terms here of interest (i.e., cluster-terms, Achinstein's A-terms: copper, metal, oxygen, electron, tiger, fish, etc)) with terms such as "grandfather," "square knot," "shoat" --which have properties linguistically associated with them. Such terms refer to objects, relationships, or phenomena which not only are relatively stable and recurrent in human experience, but upon which the advances of knowledge have no or little potential effect. But advances in knowledge has great potential for realignment of our conceptions of the cluster-terms
mentioned above (cf: history of science). We need associated properties for the latter terms, but not linguistically associated properties. It seems to me that these associations are at no time mere features of language—except insofar as knowledge of the world is embodied in language use. Consider for example the cluster of properties associated with the term "copper": color, malleability, conductivity of heat and electricity, specific gravity, density, atomic number, etc. Is any (cluster) of these properties per se linguistically associated with copper? Instead, it seems to me that individually and collectively, i.e., as a cluster, they embody our knowledge of copper, and that, in turn, this knowledge may be embodied or incorporated into our language. In other words, the association of a cluster of properties with a term is first and foremost empirical, and has to do with the world and our knowledge of it, even if eventually it becomes linguistic (cf: Quine, 1953; 1969).

Concisely put, the basic claims (N1-N4) of the new cluster theory can be formulated as follows: [All new notions and principles will be explained in a moment].

N1: A cluster of correlated and assertible properties associated (on empirical grounds) with a term is shaped, influenced or constrained by a multivariate background context, relative to a given community.
N2: (a) The 'meaning' (as use) of a term is best characterized in terms of a cluster of correlated and assertible properties associated with it, mostly on empirical grounds.

(b) A cluster-term (and therefore the 'meaning' of a term) is to be understood in relation to other cluster-terms, in the context of a cluster-network (lexical or semantical network) which has a structure, exhibits clustering-levels (or hierarchies of levels) and specific relations (of various kinds) between the various cluster-terms composing it, and may thus constitute contrast clusters (contrast sets). 9

N3: The cluster of correlated and assertible properties associated with a term and which best serves to characterize its 'meaning' over-determines (or is under-determined by) the referent of the term.

N4: An individual is said to know how to use a term if s/he can provide a cluster of the correlated and assertible properties that s/he believes (perhaps wrongly) is associated with that term, or if s/he can defer in her use of the term to relevant experts or authorities. 10

Several of the distinctive features and principles of the new cluster theory must now emerge more clearly. It can be represented schematically as follows:

Multivariate
Background
Context

\[\text{Term: 'Meaning' \quad \rightarrow \quad \text{Referent(s)} }\]

\[\quad \quad \quad \quad \quad \quad \quad \text{as Use } \quad \quad \rightarrow \quad \quad \text{[A cluster of correlated and]}
\]

\[\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{[assertible properties]}\]

\[\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{[associated with term]}\]

\[\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{[on empirical grounds]}\]

\[\quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \text{[constrains]}\]

As proposition N2 (a) above indicates, the cluster of properties associated with a term does not determine or
contribute its meaning --even though such a term does have meaning (cf: Achinstein, 1968, p.31; Paul Ziff, 1960, chap.5). Yet, we shall be concerned with the cluster of properties associated with a term and which best serves to characterize its 'meaning,' because we can reasonably assume that all we can say or assert about a given entity (referred to by a term) is given by such a cluster of properties ascribed to that entity (Shapere, 1989). Relative to a given community and its members, we shall thus be concerned more specifically with the reasons (mostly empirical) for ascribing a cluster of properties to a given entity within a "multivariate background context." Furthermore, since our concern is with the problem of comparing concepts in different theoretical background contexts, it makes sense to map out continuities and changes having to do with ascriptions of properties or clusters thereof. In this context, what is being talked about (continuity of reference) could be established by the fact that there are reasons for changing the cluster of properties ascribed to a given entity.1

Proposition N2 (b) also suggests that our theory is constrained by some form of "local holism" in that the 'meaning' of a cluster-term is always to be understood in relations to other cluster-terms, within the context of a "cluster network." This is by the way the same idea that Kuhn expresses in his 1980 Lectures using the notion of
"lexicon" or "lexical field." I take it that Grandy also gets to the same conclusion with the notions of "semantic fields" and "contrast sets" (see 1986, 1990 for details). It is also worth adding that this idea is explored -- though in a different direction -- by recent work in the cognitive sciences, where there is much talk about representation -- models in terms of semantic networks, frames, structures, schemas, etc.

Finally, proposition N2 (a) makes use of two principles (that of correlation, and that of assertibility). I will consider briefly the first one and leave the second until my discussion of N3. The principle of correlation in use in the present cluster theory is in fact double: (i) it serves to express the idea that the properties constituting a cluster do so by virtue of some cor-relation(s) holding between these properties; (ii) it serves to express the idea that the 'meaning' of a cluster-term (i.e., which is characterized in terms of a cluster of properties) is to be taken in relations to other cluster-terms within a "cluster-network" -- which can and should be further specified as part of the 'meaning' of a cluster-term (cf: Grandy, 1986, p.274 for a similar point). Concerning (i), I think that the idea is made plausible because, whether ascribed or discovered, properties rarely, if at all, 'come one-by-one' but rather as clusters; they form clusters by
virtue of the relations holding between the properties. [Would it unreasonable to assume that 'The world is (or appears to be) clustered'? cf: Gasking, 1960; Strawson, 1960; Sprigge, 1970; for some insights into a cluster-based metaphysics]. Concerning (ii), its plausibility derives from a few basic linguistic facts: terms are never (or almost never) introduced into a language as a single term, but as a cluster of terms, whereby each cluster-term contributes to the 'meaning' of the others; terms are (almost) never learned or acquired one-by-one, but as a cluster of inter-related terms (cf: Kuhn's examples of swans/geese/ducks; force/weight/mass; phlogiston/principle/compound/element/etc, 1974; 1982; 1988). Similarly, one could say that the following cluster of terms must be learned together: copper/gold/iron/metal/ore/mineral/etc.12

Considering proposition N3, we should recall first that there are reasons (mostly empirical) for associating a cluster of properties with a term or for ascribing a cluster of properties to a given entity or type of entity. In this context, the principle of assertibility does what it is supposed to do: it serves to characterize what can be said and what cannot be said, what reasons can or cannot be given, what knowledge-claims or existence-claims (or referential-claims) can or cannot be made. As such, it serves as both an epistemological and ontological constraint, relative to a multivariate background context,
which can only be delineated in turn relative to a given community and its members. It is thus a fairly 'neutral' and context-sensitive principle, which has the further advantage of not involving the kind of vicious circularity plaguing Achinstein's principle of semantical relevance. In addition, it is in accord with the open-ended nature of the process of meaning-making in science: as is to be expected, the conditions of assertibility will vary with advances in scientific knowledge.

Note that the "background context" is multivariate (as N1 states) in the sense that it may contain as many and as different elements as may be possible and relevant for a given domain of scientific investigation and research: conceptual and other elements from several different theories, beliefs (metaphysical and otherwise), knowledge and existence-claims, facts, models, experimental/instrumental data, methodologies, etc; it is much more "hodge-podge," less "monolithic" and less "deterministic" (as in Shapere, 1984) than has been assumed by proponents of traditional contextual theories of meaning. Its role, "context-specific," is to shape, influence, or simply constrain what can be asked, asserted or even researched --from a scientific point of view.13

Proposition N3 also reveals the use of the principle of over-/under-determination. As I already pointed, it serves
essentially to capture Frege’s idea that different senses may still pick out the same referent. Similarly, in the new cluster theory, different clusters of properties may still pick the same referent, by virtue of the fact that a cluster of properties over-determines (or is under-determined by) the referent of the term — in the sense explained earlier. A cluster of properties may change as a result of theory change and yet still pick out the same referent that was picked out in a previous theory (by a different cluster of properties). This principle is important insofar as it enables us to account for (trans-theoretical) changes in ‘cluster of properties’ without (trans-theoretical) reference-changes.14 Let us recall however that, in contrast with a reference theory like Kripke’s or Putnam’s, which tend to assume “rigid or fixed reference” (i.e., an essentialism of reference), the cluster theory here in question wants to account for the fact that both cluster of properties and reference may change over time. In this sense, it assumes neither an essentialism of ‘meaning’ (or meaning), nor an essentialism of reference (cf: Shapere, 1982; Kuhn, 1990).

Turning finally to proposition N4, we should note immediately that since in the new cluster theory the locus of meaning is the community (not the individual), epistemic access is not required on the part of the individual, even though some minimal or community-based linguistic competence
must be assumed (i.e., at least capacity to defer). Thus the individual (may be but) need not be able to provide or know the cluster of correlated and assertible properties associated with a term, and still succeed in using the term to refer to what it is supposed to refer. In fact, the individual may even be wrong in associating a given cluster of properties with a term, and still succeed in referring to what the term was supposed to refer to. In these cases (and possibly others), I argue that the principle of deference comes into play. Before explaining briefly, how this is so, I would like to point out that such a principle is in fact a corollary of Putnam's [and Kitcher's] principle of the division of linguistic and cognitive labor (Part II, Chap.2).

According to Putnam, every linguistic community exemplifies this principle, and that is, it possesses (at least some) terms whose associated 'criteria' are known only to a subset of the speakers who acquire the terms, and whose use by the other speakers is 'parasitic' or depend upon a structured, albeit tacit, cooperation between them and the speakers in the relevant subsets. Thus, by way of some sort of principle of deference, the way of recognizing possessed by the expert speakers is also, through them, somehow possessed by the linguistic community as a collective body even though it is not possessed by each individual member of
the community; in this way, the most recherche fact or piece of information about gold, copper, lead, water, fish, electricity, quarks, electromagnetic fields, etc., may become part of the socially and community-based meaning of the term while being and remaining unknown or inaccessible to almost all speakers who acquire the term. Even within a scientific community, which is in a sense already a "community of experts," and in which one expect individual members to have a relatively high level of (disciplinary) linguistic competence and epistemic access, and thus to be able to provide and to know the cluster of properties associated with a term, I think that a principle of deference still applies. For example, in such cases as when a scientist (or group of scientists) defers in his use of a term to the use of the scientist who originally introduced it into the scientific language. There may even be deference of the sort mentioned here between two sub-groups of a scientific community or between two scientific communities, e.g., French chemists of 18th century deferring to Lavoisier and his collaborators in their use of the term 'oxygen'; British chemists deferring to French chemists and vice versa.16

In addition to the principles presented so far, the new cluster theory will also appeal to the principle aptly dubbed the principle of humanity by Grandy (1973), [rather than to a principle of charity].17 According to Grandy,
such a principle requires us to impute to the speaker whom we are trying to understand, to translate or interpret a "pattern of relations among beliefs, desires, and the world [which is] as similar as possible to ours." (1973, p.443). Such a principle would function as a constraint because it compels us to recognize that the intentions of the speaker on the occasion of an utterance or statement play an important role, and that they ought to be identified only in relation to, and together with what we know about the speaker's or writer's environment and behavior, and our current best understanding of the pattern of relations among beliefs and desires, on the one hand, and the world, on the other, which is shared by human beings. And so our task would be to use our understanding of the patterns of relations between mental states and the world which is common to humanity, together with the available data about the speaker's or writer's environment and behavior, to identify the intentions which were operative on the occasion of utterance and thus construct an interpretation of the terms used. This is what historians and philosophers of science should be doing, and what the most seasoned among them have already been doing.18

By way of partial summary, it might be helpful to note that the new cluster theory advocated here is based on a number of related principles, which I shall divide
artificially into structural principles and interpretative principles as follows:

<table>
<thead>
<tr>
<th>Structural Principles</th>
<th>Interpretative Principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principles of Correlation</td>
<td>Principle of Deference</td>
</tr>
<tr>
<td>Principles of Over-/</td>
<td>Principle of the Division of Linguistic and Cognitive</td>
</tr>
<tr>
<td>Under-determination</td>
<td></td>
</tr>
<tr>
<td>Principle of Assertibility</td>
<td>Principle of Humanity</td>
</tr>
</tbody>
</table>

Briefly, let me say that the structural principles serve to characterize the structure of the cluster of properties in question, i.e., its composition or constituency, the relations holding between elements, items or properties constituting the cluster, and finally, the nature of the relationship between the properties constituting the cluster and what this cluster pick out as referent. As for the interpretative principles, they serve to characterize the constraints which can be brought to bear on (i) the semantical aspect of the use of a given term by a given individual or group of individuals, and on (ii) our "hermeneutical" efforts to capture and understand such a use at any given point in history.

Though I have distinguished these principles, it must be stressed that they are in fact related. Thus, the three structural principles are related to one another: for
example, properties which are assertible are very likely to be correlated in some fashion or other; and taken together, as cluster, they might over-determine a referent, or, to put it differently, the referent might be under-determined by the cluster of correlated and assertible properties. Concerning the interpretative principles, as I already pointed out, the principle of deference is a corollary of the principle of the division of linguistic and cognitive labor. In some sense, one could argue further that the principles of deference and of the division of linguistic and cognitive labor are both corollaries of the much broader and general principle of humanity (Grandy's point, personal communication, October 10, 1990). Finally, it must be noted that structural and interpretative principles are also related in some fashion: the latter constrain the former, or else, the latter are somehow implicit in the former.

Apart from exemplifying and making use of these various principles, the cluster theory envisaged here will satisfy the remaining adequacy requirements formulated by Shapere and Nersessian. It will thus be able to solve both the representational and developmental problems discussed by Nersessian, and, using Shapere's CORC device (chain-of-reasoning connections), to trace and map out across different background contexts the changes in the use of scientific terms. However, the picture that will emerge
will be different; it will be a richer model than those proposed by Shapere and Nersessian. Such a model, I argue, will enable us to account more fully for the process of meaning-making in science because it is more nuanced. It will enable us to account for the various kinds and degrees of 'meaning' and 'reference' changes (or rather conceptual changes) in science. [Contrast with Shapere's and Nersessian's account, whose concern is mainly with changes in 'meanings']. It will also account for that which has so far escaped most philosophical accounts, namely, radical conceptual change without discontinuity, and even, radical conceptual change as a result of (or simply within) a broader framework of continuous conceptual change (see forthcoming case-study of the "chemical revolution"; Part IV, chap.2, for a more vivid illustration of this point).20
B. Representational Problems

Beginning with the representational problems, we will assume a proper distinction between a term-type (general term) and instances or tokens of a term-type. There is only a token of a type, and to a type may correspond one or several tokens, or none at all. This distinction is important because it enables us to "context-sensitize" our theory, and to consider particular uses of a term by a given scientist or group of scientists in different contexts, and over time.

The 'meaning' of a term-token can be represented at some point in time, for a given scientific community, as a cluster of correlated and assertible properties -- given a multivariate background context. Depending on the kind of term, such a cluster might include information about: (i) physical/superficial/manifest properties (e.g., color, size, shape, smell, taste, etc), (ii) ontological status, (iii) function, (iv) behavior, (v) composition, (vi) structure (manifest/underlying; mathematical or otherwise), (vii) causal effects, (viii) similarity/difference relations to other entities, (ix) other kinds of considerations (e.g., associations with persons, places, events, etc). Note that unlike Nersessian, I stress the correlation between properties; furthermore, in addition to the kind of properties cited by Nersessian, I think, we should include (i) (v) (vi) (viii) and (ix) in order to be able to account
for a wider range of scientific terms (e.g., like "copper," or like "field" or "phlogiston"), and for the development from a commonsense or proto-scientific use of a term to its scientific use. Also, in accord with what I said earlier, the cluster of properties (cp) associated with a term-token (at some point in time) over-determines (and is underdetermined by) the referent(s) of that term-token.

\[ \text{Term-token: cluster of properties (cp) \text{\rightarrow} referent(s)} \]

[correlated/assertible] \text{under-determines}

Just as there are reasons (to be specified in relation to a given background context) for associating a certain cluster of properties with a term-token, there are reasons for changing the cluster of properties. As a term is used across different contexts, new properties may be added to the cluster of properties previously associated with that term (expansions), old ones may be deleted (contractions), or the cluster of properties may be restructured or reconfigured to various degrees. As these changes may take place over time in these various ways, we may end up sometimes with an altogether new cluster of properties (serving to pick out a new or different referent).

As for a term-type, it can be represented by a Cluster of properties (Cp), i.e., a Cluster of CORC-related clusters of properties. This would give us a summary (over time) of
the various clusters of properties which have been associated with tokens of the term-type in question. Corresponding to the Cluster of properties associated with each term-type, there is a "field of referential possibilities" (FRP hereafter).

**Term-type:** Cluster of properties (C)p \( \rightarrow \) Field of
\[ \begin{align*}
\mid & \quad \text{Referential} \\
\mid & \quad \text{Possibilities} \\
\end{align*} \]

[CoRC-related clusters (FRP).
[of correlated/assertible properties]

The notion of FRP is similar to Kitcher's "reference potential" (1978) in the sense that it serves here as well to characterize in a more context-sensitive way the changes in reference, what can or cannot be a referent, for a given term given a particular background context. In Kitcher's words, it is a "compendium of the ways (actual/possible) in which the referents of tokens of a term are fixed for members of a community" (1982, p.340; italics added). But it is different in one important respect: the theoretical motivation behind it. Kitcher introduces his notion in support of a reference approach in order to dispense presumably with the kind of (descriptivist) approach that I am here advocating under the name of cluster theory.

Though the representational model that I advocate as part of the new cluster theory is similar in some respects to Shapere's and Nersessian's, it is significantly different in others. [Compare Tables 1 and 2-3 below]). Like the
Shapere/Nersessian model [Table 1], it would show whether or
not, in what ways, and for what reasons the clusters of
properties (cp’s) associated with tokens of a term-type have
changed over time. It would also reveal the CORC’s between
the various tokens of a term-type, between the various
phases of development of the concept corresponding to it,
the earlier and later forms of the concept. In addition
however, and consistent with one of its major premises, the
Cluster model would also bring out and enable us to chart
the changes in the referent(s) associated with different
tokens of a term-type [Table 2]. It would enable us to map
out the various shifts, expansions, contractions or
restructurations in the [Cp] and [FRP] of the term-type in
question [Table 3]. The major premise referred to above is
that since the relation holding between ‘meanings’ (cluster
of properties) and ‘reference’ is neither strict, nor
all-deterministic, in other words, since we have introduced
some “play” between these notions, we must be concerned with
characterizations of both ‘meanings’ and ‘reference’ at some
point in time and over time, as they may both change.

Strictly speaking, the Shapere/Nersessian model can be
illustrated as in Table 1 (see Part III, Chapters 1 and 2). When
charitably interpreted and improved in light of the
major premise discussed above, it should look as in Table 2,
and should apply in some cases (e.g., ‘electron,’ ‘field,’
and even 'copper'). As for the Cluster model, it suggests a richer and more complex picture as in Table 3 which applies in addition to other cases (e.g., 'phlogiston'/'oxygen', 'ether', 'caloric heat')—including the one case which shall interest us. It should be obvious, when all implications are drawn, that the Shapere/Nersessian model is in fact a particular (and simpler) case of the Cluster model advocated here.
**Table 1:** Shapere/Nersessian Model

<table>
<thead>
<tr>
<th>Tokens of a Term-Type &quot;T&quot;</th>
<th>Background Context</th>
<th>Properties Ascribed (none essential)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>BC1</td>
<td>ABCD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>BC2</td>
<td>ABC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>BC3</td>
<td>ABCF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>BC4</td>
<td>ABGF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td>BC5</td>
<td>HKGF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2:** Shapere/Nersessian Model [Improved Version]

<table>
<thead>
<tr>
<th>Tokens of Term-Type &quot;T&quot;</th>
<th>Background Context</th>
<th>Cluster of Properties</th>
<th>Referent(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>BC1</td>
<td>ABCDE</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>BC2</td>
<td>ABCD</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>BC3</td>
<td>ABCEF</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>BC4</td>
<td>ABGHL</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T5</td>
<td>BC5</td>
<td>GHKL</td>
<td>R</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: New Model [Cluster theory]

<table>
<thead>
<tr>
<th>Tokens of Term-type</th>
<th>Background Context</th>
<th>Cluster of Properties</th>
<th>Referent(s)/FRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 BC1</td>
<td>ABCDE</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2 BC2</td>
<td>ABCD</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3 BC3</td>
<td>ABDFG</td>
<td>P, [A ?]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4 BC4</td>
<td>BDFGH</td>
<td>P, [A ?]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* * * * * * * * * * * *

O1 BC5 DFGH [P, X ?]

O2 BC6 FGHKL X, [P ?]

O3 BC7 FGKL X, [H ?]

O4 [Y1?] BC8 GHKL M [?], X [H ?]

*Vertical lines indicate CORC's (chain-of-reasoning connections). * * * * * indicates revolutionary scientific change. Note: this model is intended to reflect at least in part the complexities involved in the case-study of the chemical revolution (Part IV, Chapter 2).

Though these models clearly fail to exhibit all the complexities of scientific change, they give us some idea of what may be involved in at least two different kinds of cases. In Table 1, we can the trace the changes
(expansions, contractions, and restructurations) in the cluster of properties associated with a term 'T' and used in different contexts to refer to a particular entity (R): though the cluster of properties associated with the use of the term 'T' at different times has changed, the referent (R), and consequently, the FRP has remained the same [Table 2]. In Table 3, on the other hand, we can trace the changes endured by the cluster of properties associated with a term 'F' and by its corresponding FRP; as a result of radical and revolutionary change in BC [BC4 ——> BC5], indicated on the diagram by the starred horizontal line, the term 'F' has been dropped in favor of 'O', and the entity (or entities) referred to by 'F' has (have) also been progressively abandoned and replaced ultimately by another entity (X). In this case, we can trace the changes (expansions, contractions, restructurations) in both the Cluster of properties and FRP associated with a given term. Thus, despite radical change, one can see how it took place continuously, i.e., within a larger background of continuity.

More specifically, one may come to discover for example that, prior to the revolutionary change (indicated by ** *), 'F' 'partially referred' to (P) in some context, and possibly to (A) as well, in another context; one may discover in post-revolutionary developments that (P) is in
fact (X), and that (A) (or what was thought to be (A)) is in fact (H), and subsequently, the term 'O' rather than 'F' may be used to refer to the new entity (X) and another term (e.g., Y1) may be used to refer to (H).

Incidentally, it is worth noting that this description reflects to a large extent the actual developments which took place in the transition from phlogiston theory to oxygen theory in the history of chemistry (see Chapter 2 for case-study). Supposing that the phlogistonists believed that the term 'phlogiston' designated that substance which was emitted during combustion, one could then say that 'F' designated (P). But as the historical record suggests, they also believed (at least for a time) that inflammable air (A) which they had isolated was phlogiston. Once they had made this identification, they went on to record the properties of inflammable air (which was in fact hydrogen) using the term "phlogiston" (See Priestley's letters [1782-4], in Schofield, 1966, esp. letters #93, 95, 98, 114, 115; see also Kitcher, 1978, pp.533-4, note 17).

As I suggested earlier [at the beginning of my discussion of representational problems], the kind of cluster of properties associated with a term will vary with the kind of term that we are dealing with - e.g., substantive/natural kind, qualitative, dispositional, quantitative, etc. However, for each particular case, we can attempt to
determine which kind of cluster of correlated properties is assertible relative to a given background context for a given scientific community—at some point in time and over time. For example, we should be able to say something about whether these properties are of one sort or another, their relations to one another, their salience or centrality, their degree of assertibility, and possibly even about the ways in which they—as a cluster—over-determine the referent(s) of the term associated with them. This can only be accomplished however only if we always begin by "looking at what scientists do and say" in situations where new concepts are formed and new terms introduced; at the reasons they give for associating a given cluster of properties with a term in a certain way. Only a historical/cognitive account of such practices by actual scientists will do.24

Such an account should enable us to map out the changes (if any) in the "Cluster of properties" and "field of referential possibilities" associated with a term-type. Since the reference of a term-type is not fixed once and for all, and since different tokens of a term-type can refer to different entities, what we need is a context-sensitive account (cf: Kitcher, 1978). I submit that the new cluster constitutes such an account, as I have suggested throughout. It should tell us what the referent of a token-term is (if anything at all, i.e., if there is anything for it to refer
to); whether a token-term picks out a new referent (or several); whether the previous referent of a token-term is abandoned -- and the term along with it; or whether in addition to the previous referent, a new entity is also considered to be a possible referent, etc [see Table 3 for nuances along these lines]. In other words, it should be able to say whether the Cφ and FRP corresponding to the term-type has expanded, contracted, or been restructured. Our cognitive/historical account of the production and use of scientific terms by scientists, of their meaning-making and referential activities should be constrained, as I have already argued, by the principle of humanity.

Suppose we want to formulate such a referential account for the language of phlogiston theory or oxygen theory, or any other scientific theory. It is reasonable to assume that most terms in this theory will be context-sensitive; furthermore, some terms will be "homogeneous" in that their FRP is determined by one referent (one entity), while others (perhaps most) will be "heterogeneous" in that their FRP is determined by two or more referents (two or more entities). [See Table 3]. These terms can properly be called "theory-laden" (Kitcher, 1978). Bearing this in mind, and the principle of humanity (or some other general principle), our task consists then in formulating "empirical hypotheses" about the referents of token-terms of that type which will
satisfactorily explain the behavior (linguistic and otherwise) of a proponent of the theory. 25

An obvious and immediate objection would be to point out that in some circumstances two or more hypotheses about the referent(s) of a scientist's tokens may provide equally good account of his/her behavior (linguistic and otherwise). In such circumstances, the principle invoked will not help presumably decide between hypotheses. An obvious and immediate response is that even if the principle invoked (and any other principles) may leave our ascriptions of reference undetermined or indeterminate (to some extent, in some cases), this does not support the extreme and radical theses of incommensurability and incomparability between scientific theories.

In fact, as we probe deeper into these issues having to do with the referential activities of past (or other) scientists, we can see that different situations may obtain --none of which constituting a serious threat to commensurability and comparability in science.

We may be able to find and specify type-type correlations between terms of theory T1 and terms of theory T2, but, as Kitcher has shown, this is very unlikely and to say the least unrealistic. Instead the following situations are more or less likely. (a) We may be able to specify the referent(s) of each token-term used by a proponent of T1;
(b) we may be able to specify the referent(s) of some
token-terms used by a proponent of T1; (c) we may be able to
specify some of the referent(s) of each token-term used by a
proponent of T1; (d) we may be able to specify some of the
referent(s) of some of the token-terms used by a proponent
of T1; (e) we may not be able to specify the referents of
some token-terms of T1; however, we may be able to specify a
cluster of entities such that the referent of any token of
that type belongs to the cluster in question; and finally,
(f) for some terms-type, we may not even be able to specify
a cluster of entities such that the referent of any token of
that type belongs to the cluster in question.

According the cluster theory here advocated, none of
these situations, not even (f), threatens the rationality
and objectivity of science, the commensurability and
comparability of scientific theories. Situation (a) is
clearly the least problematic: it represents what we may
call "ideal context-sensitive determination of reference."
Situation (b) involves "partial determination of reference"
(contrast with Field's notion of "partial reference" on the
intuitiveness of each notion); it is also unproblematic for
the most part. Situation (c) may be considered to be a
situation involving "partial context-sensitive determination
of reference." Situation (d) in turn may be characterized
as one involving "partial determination of partial
reference."

As for situation (e), it may be characterized as the case of "under-determination of reference." This case is already begins to announce a worst possible scenario. We must consider a number of different ways (different theoretical hypotheses) of formulating (parts of) of the past language or theory under consideration. But one should realize that this does not preclude the possibility of commensurability and comparability. Finally, situation (f) may raise the specter of extreme and radical incommensurability and incomparability, and it is, as Kitcher rightly noted (1978), the one which is often mentioned in support of this thesis. The general claim is that when attending to the language of a past scientific theory, across a revolutionary divide, we are confronted with situations of type (f), which make this language or theory incommensurable and therefore incomparable with a modern or contemporary language. But, even in such situations, the strong theses of incommensurability and incomparability are not justified.

If such situations may be characterized as situations of "partial inscrutability of reference" or "partial indeterminacy of reference," then what may arise as a result can only be "local incommensurability." And there is no reasons to believe that such occurrences will fully
undermined the rationality and objectivity of science, for there will always be much more that is shared or agreed upon by proponents of two different theories, even when they are separated by a revolutionary divide. Consequently, they will be able for the most part to formulate their disagreements [see case-study of the chemical revolution: Kirwan's debate with Lavoisier and his collaborators; Cavendish's "manual of translation" of his theory into Lavoisier's; Priestley's arguments against the "anti-phlogistinians"].26

In the preceding discussion, my aim was not to pretend that we can always decide which cluster of properties is (best) associated with a given token-term, what the referent(s) of that term is (are), nor to claim that we can always identify or agree upon the Cluster of properties of a term-type, and its corresponding FRP. However, I still want to suggest that we can very often make such determinations, and we can thus map out the various kinds and degrees of meaning and reference changes in science. Claims about such changes for a given term-type are best understood as changes in the Cluster of properties and, concomitantly, as changes in the FRP corresponding to it. These notions may be seen as part of the complex apparatus that the new cluster theory proposes in order to exhibit the fine-structure of the history of science, of the process of meaning-making in
science.

In order however to account more satisfactorily for the complexities of scientific change, in accord with one of the main tenets of cluster theory as envisioned here, we must also keep in mind that that the 'meaning' of a cluster-term (associated with a cluster of properties is to be understood only in relations to other cluster-terms (as part of a Cluster of terms) within what we might call a "Cluster Network" --another name for "conceptual scheme," "semantic or lexical field" [?]. It must be stressed that the new cluster theory differs radically in this respect from the Shapere/Nersessian account.

A Cluster Network might be understood in terms of cluster-points (each corresponding to a cluster-term (or concept)), related in various ways to one another. These relations holding between cluster-terms in a given Cluster Network might include (see Kuhn, 1982; 1988; Grandy, 1986, pp.275ff; 1990, pp.1-15; Thagard, 1990, pp.184-6):27

(i) a kind-similarity/difference relation
(ii) an exemplification/instantiation-relation
(iii) a typical/general (or universal [?]) relation
(iv) a property-predication relation
(v) a containment (inclusion)-relation
(vi) a part/whole relation
(vii) a permutation relation
(viii) a transitive relation
(ix) higher order relations
(x) other relations [?].

They structure and organize the Cluster Network (and by the same token the Cluster Space corresponding to it) by
constituting contrast-clusters and furthermore hierarchies or clustering-levels. These clustering-levels can be of various sorts depending on the kind of relation(s) involved, e.g., (i), (ii), (v) or (vi). Cognitive scientists have repeatedly underlined the importance of these kinds of clustering-levels in that they serve to organize and structure the mental lexicon (Kuhn, 1988). But, as Thagard suggests (1990, p.185-6), they might be also useful for understanding some of the specific mechanisms underlying radical or revolutionary conceptual change. To get an idea of what a Cluster Network looks like, contemplate the following fragment for minerals:
A Fragment of the Cluster Network for Minerals

Assuming that a Cluster Network consists of cluster-points and relations such as the ones just cited holding between them, how shall we then characterize conceptual or meaning change? We could say that it consists of reason-based expansions, contractions, and/or restructurizations of the Cluster network; or else, of additions, deletions, and/or replacements/reorganizations of cluster-points and/or relations in the network in question. Given the nature of the Cluster Network, changes take place "holistically" or else in a "locally holistic manner," i.e., in clusters
rather than one-by-one. Thus, the addition (or deletion) of a cluster-point may be accompanied by the additions (or deletions) of various relations and other cluster-points. However, it is important to note, all additions or deletions are not of the same and equal consequence. The additions of new cluster-points as well as new relations such as (i), (iii) and possibly (vi) [mentioned above] --where new cluster-points (terms or concepts) and new relations replace those of a previous Cluster Network, will be considered as involving the most dramatic changes. Replacement of a Cluster Network can be distinguished from simple additions or deletions: in the latter cases, some of the previous relations still hold, indicating thus that the new cluster-points (terms or concepts) and new relations have a role and place in the new Cluster Network which are similar to those which existed in the old Network. Thus, although a scientific revolution can involve dramatic replacement of a significant part of a Cluster Network, there is nevertheless continuity in the sense that some relations are always maintained. 28 I submit however that the dramatic and radical changes brought about by a scientific revolution will be most apparent at the clustering-levels defined by relations (i), (ii), (iii), (vi) and possibly (ix) [see Thagard, 1990, p.187]. For, changes involving these kinds of relations will usually entail a restructuration or
reorganization of the Cluster Network which is significantly different from those involved in simple additions or deletions to an established Cluster Network.
C. Developmental Problems

So far I have concerned myself mainly with representational problems that any adequate theory of conceptual change and of the process of meaning-making in science must solve; but, as I indicated, there are developmental problems which must be addressed as well before we can have such a theory. However, one should keep in mind that the specification and eventual resolution of these developmental problems are closely tied to the specification and resolution of the representational problems. In passing, it must be noted that one of the advantages of the cluster-representation is that it suggests a diachronic dimension even while being applied synchronically (Part I, chapter 3). When talking about developmental problems, we are not just concerned with the historical dimension of meaning-making in science, i.e., with how conceptual networks change over time. This concern, I will assume, has been dealt with already to a sufficient extent. Rather what we are talking about are (i) the cognitive mechanisms and processes by which scientists discover and build new conceptual schemes (or new Cluster Networks). [From here on, I will talk about conceptual (Cluster) networks]. But we are also interested in (ii) the mechanisms and processes by which a new conceptual network replaces another; (iii) the mechanisms and processes by which other members of the scientific community can acquire,
learn and come to accept (and even pursue) the new conceptual network and the research program that it opens up, and subsequently, (iv) the mechanisms and processes by which members of the scientific community (other than the scientist-discoverer) come to replace the old conceptual network with the new. Any adequate theory of the process of meaning-making in science must be able to account for these problems, if it is to be able to account for revolutionary conceptual changes (compare and contrast with Nersessian, 1987;1989; see Thagard, 1990. p.197).29 In passing, Nersessian's account confines itself to only (i) and (ii).

Traditional accounts of scientific conceptual change can be roughly subdivided into what we might call (i) progressive theories, or accretions or incremental theories, as they are sometimes (wrongly) called,30 and (ii) gestalt theories (Thagard, 1990, p.194).31 The former holds that a new conceptual network develops by progressive modifications, combinations, differentiations, refinements, accretions, increments or simply by additions of new ones, and possibly by deletions of existing cluster-points and relations, (Nersessian, 1987; Rumelhart, 1980, pp.52–54). The underlying assumption is that the new conceptual network emerges somehow within a gradual, continuous and cumulative process, and somehow extends or modifies the previous one progressively. The latter theories on the other hand hold
that a new conceptual network emerges abruptly and in a radical discontinuity, as a result of "gestalt switches" of the kind that occur in perceptual phenomena (e.g., duck/rabbit; Necker cube), and completely replaces the previous one. The underlying assumption is that this all happens in a discontinuous and disruptive manner which involves reorganizations, additions, and replacements. This is presumably what happens as a result of a scientific revolution, and which leads to the incommensurability problem (Kuhn, 1970; 1988). 32

Neither theory, taken alone, seems to be able to account adequately and fully for the kind of conceptual change which takes place as a result of a scientific revolution (see Thagard, 1990, p.196). For example, progressive theories seem inadequate to account for the kind of radical and significant degree of restructurations or reorganizations as well additions and replacements which occur typically in such a case. On the other hand, gestalt theories seem unable to accommodate the fact that there is continuity in conceptual change, or else, despite conceptual change. These theories gain however a certain plausibility when one focuses on two extremes stages in scientific development, and attempt to relate two different conceptual networks corresponding respectively to each of these stages (e.g., Newton//Einstein; Faraday//Einstein; Stahl/Priestley/
But then they neglect all the intermediate, small steps between these extremes which have led progressively to the formulation of a new conceptual network often on the basis of an old and established one (Shapere, 1984; Nersessian, 1987; Thagard, 1990, p. 195-6). Furthermore, both kinds of theories seem unable to account fully for conceptual change, not just with the individual scientist, but among members of the scientific community who reject the old conceptual network and accept a new one. This is a significant failure since a scientific revolution occurs only when a scientific community as a group adopts a new conceptual network (contrast with Nersessian, 1987; see Thagard, 1990, p. 197).

Therefore, what we need is an account which avoids the limitations of both progressive and gestalt theories, and which accounts for conceptual change, not just in an individual scientist, but in the scientific community at large. A cluster theory such as the one that I envisage can make a significant contribution in this respect. Insofar as the mechanisms and processes here of interest have to do with the cognitive and psychological experiences of individual scientists, and insofar as these experiences (can) have broader social and historical consequences on the (scientific) community at large, we might benefit in our endeavor not just from recent contributions in cognitive
science (cognitive psychology, artificial intelligence)

--as Nersessian suggests, but also from cognitive history and sociology of science [see for example Tweney, 1981; 1985; Gibson, 1985; Thagard, 1988, 1990; Giere, 1988; Simon et al, 1987; Barnes, 1980; McMullin, 1988; Donovan et al, 1988].

To be more explicit, let us assume a distinction between "context of discovery" and "context of justification." An adequate theory of conceptual change (and of the process of meaning-making) in science must account for the following developmental problems in both contexts:

**Context of discovery:**
(1) Cognitive mechanisms and processes by which an individual scientist develops a new conceptual network.
(2) Cognitive mechanisms and processes by which an individual scientist comes to reject the old and replace it with the new conceptual network.

**Context of Justification:**
(3) Mechanisms, methods and processes by which other scientists in the community learn or acquire a new conceptual network.
(4) Mechanisms, methods and processes by which other scientists come to reject the old and replace it with the new conceptual network.

In order to answer (1), we must (at least) consider (i) concept-formation: how new concepts and the connections between them are formed; how they are subsequently altered, modified, and replaced; (ii) forms of reasoning (induction, abduction/adduction, deduction): which particular forms of reasoning are used to what effects; in what particular
contexts; (iii) special disciplinary heuristics; (iv) experiments and generalizations (on the basis of experimental/instrumental data); and (v) hypothesis-formation (see pertinent discussions by Thagard, 1990, pp. 197-201; esp. 206).

In her answer to (1) above mentioned, Nersessian focuses mainly and only on (i) concept-formation. She argues, as we have seen (Part III, chapter 2), that the cognitive mechanisms and processes involved are best described by progressive theories. She rejects the view (Kuhn's [?]) according to which new concepts (always or often) arise by some inexplicable "gestalt" or mysterious "creative leap" which defies rational analysis, and somehow always replace former concepts—in such a way that an incommensurability of meaning develops between one period and the next in science. In her view, when concept formation is regarded and taken, as it should be, as a cognitive/historical process, it is possible to discern patterns and strategies of concept formation which are subject to some analysis; she considers in particular analogies, images, metaphors, models. A study of these patterns and strategies reveals, she claims, that the dynamics of concept formation are not those suggested by gestalt theories. Rather, they suggest that new concepts are formed for the most part by such processes as
accretions, deletions, combinations, differentiations; in short, by refinements from existing ones within an established conceptual network in a related or unrelated discipline or field of research. Nersessian is right in this respect, but, as I argued earlier, her view is limited; we need both progressive and gestalt theories.

Insofar as there will be different kinds of patterns and strategies at different phases of the process of concept formation, it may be helpful to consider if only tentatively a framework within which to discuss these different phases and the different patterns and strategies corresponding to them. One such framework might distinguish three major phases in the process of concept formation and development in science.

**Heuristic Phase**: a new way of conceptualizing some phenomena or problems is introduced primarily by borrowing ideas, analyses about (similar) phenomena or problems in a related or unrelated discipline or field of research. These are used heuristically to explore the new conceptual possibility. Typically, at least in the physical sciences, this involves the use of analogies, models, images, metaphors, etc.

**Extension Phase**: the new concept is accepted as a fruitful way of comprehending some phenomena; in which case, methods of analysis appropriate to the discipline or field of research are used to elaborate, clarify, and extend the concept in question. This may require mathematical analysis, experimentation, or simply more data collecting.

**Critical Phase**: the now relatively developed concept is subjected to rigorous critical scrutiny. Sometimes, a philosophical analysis of the foundations upon which it rests is undertaken; in the process, the concept is reassessed, further clarified or extended, and even sometimes reformulated upon new foundations.
These phases do not necessarily occur at separate times; there can be, and there usually is, interaction. Furthermore, the transition from one phase to another may taken place within the theory of one individual scientist or span several generations of scientists and a succession of theories. In general however, it seems unlikely for a concept to be taken through all three phases in the work of one individual scientist. Also, the process of articulating ways of conceptualizing some phenomena remains open-ended. Note in passing that while Nersessian concerns herself only with (i) concept-formation, Shapere's focus in this respect is on (ii) forms of reasoning.

Turning to problem (2) mentioned above, we need to keep in mind that since conceptual networks function as (more or less) integrated wholes, the replacement of an existing network by a new one is unlikely to be simply a matter of adding new cluster-points (or concepts) and relations and replacing old ones. The overthrow of a conceptual network is not accomplished just by rejecting particular cluster-points or relations, but also by challenging the entire structure and replacing by a new and well-articulated alternative. If somehow incremental theories are better suited for building and developing a new network, gestalt theories seem more appropriate to account for replacement, since they claim that a whole new conceptual network can
emerge all at once, as a result of a gestalt-like pheno-
menon. Even so, these kinds of theories do not explain how
the new conceptual network can be constructed progressively
and how exactly the replacement can occur. In particular,
they do not account for the fact even though the conceptual
networks separated by a scientific revolution are different
in many serious respects, they still (may) agree on a host
of issues (e.g., other concepts, 'observations',
experiments). In other words, they do not account for the
fact even revolutionary, radical and discontinuous
conceptual change takes place against a background of
continuity — with concepts, experiments and observations
which have some relative stability. And so, contrary to
Kuhn's suggestions that in scientific revolutions, (a) a new
conceptual network (a 'paradigm' or 'disciplinary matrix')
supplants the old all at once; that (b) a new conceptual
network brooks no rivals and that (c) "the world changes"
(1962, p.150) as a result, what we have instead is process
which can also be characterized in terms of three stages.35

Thagard proposes an account along the following lines
(1990, pp.201-4): (1) An old conceptual network — with a
certain degree of explanatory power, coherence and
problem-solving capacity — and with established connections
to a cluster of other concepts, 'observations', experiments
is (2) progressively and perhaps sporadically challenged by
a partially developed new conceptual network—which has the same connections to a cluster of other concepts, observations and experiments. As the new conceptual network is further developed and applied by the individual scientist who proposed it, s/he may experience a gestalt and become convinced that one should "look at the world" from the point of view of the new conceptual network; s/he may in other words become suddenly convinced of the greater degree of explanatory power, coherence, and problem-solving capacity of the new conceptual network, and begin arguing vehemently for its adoption in replacement of the old one; and thus, stage (3) occurs, whereby the new conceptual network replaces (for the individual scientist) the old one. However, it is important to stress, using a different metaphor, it is as if there were "ground-shifts" (in conceptual space) rather than "world changes." The new conceptual network emerges in the "background" of an established one—with similar connections to a cluster of other concepts; as it "upstages" and replaces the old one, and moves somehow into the "foreground," the old one recedes into the background but does not disappears completely—at least for a while. Interestingly, the history of science is replete with examples of scientists (e.g., Lavoisier) who have developed a new conceptual network, and therefore a new language, and who could still, when required or necessary,
resort to the old network and language. Having said that, one may still want to ask: what exactly makes an old network recede into the background and a new network move into the foreground? The short answer, as suggested above, is that it is the explanatory power and coherence, and the problem-solving capacity of the new conceptual network which becomes evident as a result of repeated and systematic applications of the network in question (Thagard, 1990, p.206).

As for problems (3) and (4), we can answer them along lines similar to those articulated above — concerning the mechanisms and processes by which individual scientists discover, develop and replace conceptual networks. Just like the latter, other scientists (members of the community) acquiring a new conceptual network have to learn about its various elements (inter-related concepts, forms of reasoning, experiments/generalizations, heuristics, hypotheses) and reinforce the network by repeated applications and uses to the point where it could move to the foreground (in conceptual space), "upstage" the old network and make it recede into the background. Often, this is a long and protracted process (which might take years or decades) and which meets with various forms of opposition and resistance for various reasons — none of which being clearly and uncontroversially determinant.

Scientists can learn about a new conceptual network and
its elements in various ways: by verbal communications, through textbooks, journals, conferences, society proceedings, direct instructions, exposure to exemplars, repeated experiments, etc. But which of the various mechanisms and processes which might contribute to reinforcing the new network for a given scientist or group of scientists are the most determinant? Though, as Kuhn as shown among other things in his 1962 work, it would be naive to think that rational arguments and logical reasonings play such a role in a clear-cut way, one cannot deny, I think, that arguments and reasonings have a significant effect in this respect—particularly arguments and reasonings attempting to show the explanatory power and coherence, the problem-solving capacity of the new conceptual network and the relative weaknesses of the rival conceptual network. However, one should add that much of the evidence from the history of science indicates that scientists are not directly and immediately convinced by arguments or by logical reasons alone; they often take years or decades before accepting a new conceptual network, and rejecting the old one; in some cases, they even waver back and forth between the old and the new conceptual network (e.g., Priestley, according to Ihde, 1980, p.84). Conceptual change takes place often at the end of a difficult, long and protracted process.
Some of the factors which might promote resistance and even opposition to a new conceptual network might include:

(i) The new conceptual network may (still) be supported by a relatively small number of experiments and limited experimental data which can be accounted for (perhaps even equally well) by the old rival conceptual network. (ii) Arguments and reasons in favor of the new conceptual network can always be met with counter-arguments and counter-reasons based on and in favor the old rival conceptual network.

(iii) The new conceptual network may still have internal unresolved problems which make it open to easy objections --particularly in light of the still appreciated explanatory power and coherence of the rival network which is still (at least partially) in place. (iv) Scientists may not have (yet) been exposed to, used or applied the new conceptual network enough to appreciate its explanatory power, its coherence, its problem-solving capacity, and its simplicity. (vi) Finally, there may be personal and other idiosyncratic factors involved which are to be determined on a case-by-case basis.37

Before turning to the case-study announced earlier, and in conclusion of this chapter, it might be helpful to further bring out the distinctive features of the new cluster theory by summarizing briefly some of the points which distinguish my account, that of the new cluster
theory, from Nersessian's, and Shapere's for that matter—despite sharing with them some general foundational assumptions.

They both explicitly reject cluster theory as a viable and plausible account of the process of meaning-making in science. But as I have argued, they do so only because they have a particular construal of cluster theory in mind, the traditional version. Instead, in this whole project, I have argued that a new version of cluster theory constitutes such a viable and plausible account of the process of meaning-making in science; it avoids or meets the traditional and common objections raised against it; it satisfies the major adequacy requirements formulated over the years for such an account.

My account assumes a different, more charitable, reading and understanding of Kuhn (1962, 1970) than Nersessian's and Shapere's. For example, incommensurability is not a false problem, as they sometimes seem to suggest in their less nuanced statements. Though it is not as insuperable and threatening as it has often been assumed, it is a real, but graded problem. In accord with the assumption of "local holism" underlying the new cluster theory, most cases of incommensurability are cases of "local incommensurability."

Nersessian and Shapere focus in their respective account on 'meanings' (at some point in time and over time); but, as I have argued, we need to concern ourselves with characterizations of both 'meanings' and 'reference' at some point in time, and as they change, over time. We also need to redefine (pace Frege's legacy) the nature of the relationship between 'meanings' (or sense) and 'reference'. While Nersessian and Shapere avoid doing so, the new cluster theory makes a proposal in this respect.

While Nersessian sketches out concept-representation in terms of (multi-component) 'vector' and 'array' or 'meaning-schema', and while Shapere proposes a concept-representation in terms of CDRC's (chain of reasoning connections), or reason-related criteria, which are to be understood as somehow part of the concept itself, my account is, I think, more fully developed, thoroughly and systematically based on the notion of Cluster and that of Field of Referential Possibilities.

Insofar as it incorporates Shapere's and Nersessian's
CORC device, it must be noted that such a device is further clarified and extended in the new cluster theory. For example, it is applied not just to trace and map out changes in the Cluster of properties (or 'meanings') associated with a term, but also to changes in the field of referential possibilities, and in the referent of the term in question.

While Nersessian is determinate and specific about what constitutes the relevant "background context" in a particular case, my account sides with Shapere in claiming that it is instead much more "hodge-podge", less determinate and deterministic than has been assumed by traditional contextual theories, and by Nersessian's account to some extent. As I pointed out earlier, it is on this basis, among others, that critics of Nersession's network account have argued that it somehow still falls prey to incommensurability, despite claims to the contrary.

In characterizing the kinds of properties which may serve in a particular case to characterize a concept, Nersessian mentions: (i) ontological status; (ii) function; (iii) behavior; (iv) structure (mathematical); (v) causality. In my account, I submit that we must include — depending on the kind of concept that we are dealing with, and the period in the history of science that we are interested in (pre-science/proto-science/science): (vi) physical/superficial/manifest properties (size, color, shape, taste, etc); (vii) structure (mathematical and otherwise, i.e., underlying structure, atomic, sub-atomic, etc); (viii) composition; (ix) other considerations (e.g., associations with places, names, events, and special properties [esoteric, medicinal, etc]); (x) relations between various properties; (xi) relations between clusters of properties, i.e., cluster-terms.

While Nersession claims that "progressive (or continuist) theories" instead of "gestalt (or discontinuist) theories" are needed in order to account plausibly for scientific concept-formation and development, I have argued that we in fact need both kinds of theories to account for the different kinds and degrees of conceptual changes which occur in science.
In order to make the case for "progressive or continuist theories" in the study of concept-formation and development, Nersessian focuses on analogies and images. But as I pointed out, and I think that Nersessian would wholeheartedly agree, we must also consider metaphors and models. In addition, I have proposed a general framework (in three phases: heuristic, extension, critical) for characterizing the process of concept-formation and development in science. Furthermore, in order to better understand the cognitive mechanisms and processes by which a scientist develops a new concept or conceptual network, we must consider not only analogies, images, metaphors, and models (Nersessian, 1987), but also forms of reasoning (Shapere, 1984), special/disciplinary/extra-disciplinary heuristics, experimental data and generalizations, and hypothesis-formation (Thagard, 1990).

Finally, in her account of conceptual change in science, Nersessian focuses only problems pertaining to the "context of discovery," namely problems (1) and (2) discussed in the section on developmental problems, and she neglects problems (3) and (4) pertaining to the "context of justification." In contrast, the new cluster theory advocated here requires that we address both kinds of problems, i.e., (1)-(4).38
Chapter 2: Case-Study of the Chemical Revolution

In this chapter, I turn to a case-study in order to show how the new cluster theory applies to a particular episode in the history of science, i.e., the "chemical revolution" in which Lavoisier's oxygen theory replaced phlogiston theory. This episode offers a very good example of radical conceptual change in science, or in I.B. Cohen's terms, "a paradigmatic example of a revolution in science" (1985, p.236). And so, our study will enable us to better understand the structure of this scientific revolution (Kuhn, SSR, 1962;1970), or better yet, the processes which have led to this revolution. But, in Shapere's view, this episode also offers an example of "perhaps the most important thing that we want a theory of scientific change to be able to do." That is, "to show how an older theory and the newer one can be compared despite the fact that the central concept in the two theories is different --despite, that is, [a] lack of common reference" or else, despite the fact that a central concept (and its corresponding entity) in an old theory is abandoned and replaced, either by a new concept (and a new sort of entity) and/or a new way of viewing a domain (1989, pp.429-30; see also Table 3, chapter 1, Part IV: new model [of cluster theory]). According to Shapere, this represents "a, if not the, crucial possibility to be explained by any adequate theory of scientific change" (1989, p.430).1 In passing, if, as Shapere again suggests,
the sort of conceptual change being focused on here is one which the causal theory of reference fails to account for, if it is of the sort which brings out still further the inadequacy of a reference account (Putnam-like), if it is also not accounted for satisfactorily by Kitcher’s improved version of Putnam’s CTR—as I have argued, and if finally the new cluster theory advocated here succeeds in providing a better and perhaps even satisfactory account of this sort of change, then the comparative advantage of the latter should be further established.

In contrast to the case-studies offered by Shapere and Nersessian, which concern themselves with a kind of conceptual change, whereby ‘meanings’ change radically while the referent remains the same, this case-study is about radical conceptual change involving both ‘meanings’ and ‘reference’ (contrast Tables 2 and 3 from Part IV, chapter 1 [B]). Furthermore, as I shall attempt to show, the application of the new cluster theory affords us the opportunity to give a far more detailed analysis of the fine-structure of the history of chemistry at this crucial juncture, of the complexities involved in the radical conceptual change which led progressively from (Becher’s/ Stahl’s) phlogiston theory to (Lavoisier’s) oxygen theory.

When Lavoisier (1743–1794) began to formulate his own views (in late 1772) about the major phenomena of interest
at the time, the dominant and established theory in chemistry was (Stahl’s) phlogiston theory. Such a theory offered presumably comprehensive and systematic explanations of the phenomena of combustion, calcination, and respiration. By 1789 however [after three other major developmental stages] chemistry was set on "a different and surer foundation," the majority of chemists had turned over to Lavoisier’s oxygen theory, which proposed a different, more coherent and simpler explanatory account of these same phenomena.

A fully developed account the chemical revolution would map and trace every stage in the development of ideas which led from phlogiston theories to Lavoisier’s most articulated oxygen theory (and even beyond). But such a task is beyond the scope of this project. Instead, I will focus on some of the most important stages in the development here in question. I will thus provide a partial and fragmentary cluster analysis of the conceptual networks involved at these different stages.

With the richer and complex apparatus of the new cluster theory in hand, I will begin by describing and characterizing the cluster conceptual network of Stahl’s (or Becher’s) phlogiston theory as well as Priestley’s, Scheele’s, Cavendish’s, and Kirwan’s later versions; I will then attempt to describe and characterize Lavoisier’s
cluster conceptual network at **four** different stages in its development, up to the formulation of his mature oxygen theory. This should enable us to map out the various kinds and degrees of conceptual changes which occurred in the development of chemistry from phlogiston to oxygen. In accord with the new cluster theory, all (or most of) the terms involved must be viewed as cluster-terms, always in relation to other cluster-terms, as part of a cluster conceptual network.3 And finally, I will give some account in terms of cluster theory of the mechanisms and processes (cognitive and otherwise) which were involved in the development of the new conceptual network, and its acceptance.

A. Representational Problems

**Phlogiston theory: Stahl’s or Becher’s?**

Depending on how one settles the *priority controversy* about the originator of phlogiston theory, one might say that it is Johann Becher (1635–1682) or Georg Stahl (1660–1734) who relied heavily on his teacher’s ideas [Partington, 1961]. For our present purposes however, we shall settle for Stahl, since his conceptual network was the most influential (see Lavoisier’s references to Stahl’s doctrine from 1772 on). It was later advocated in somewhat different and improved versions by Priestley [1733–1804], Scheele, Black, Cavendish and Kirwan --some of the most
prominent chemists of the 18th century.

To begin with, one can reconstruct Stahl’s conceptual network (at least partially) in such a way that it reveals the main cluster-points (cluster concepts or terms) that it contains and the relations assumed to be holding between them. Like Becher, Stahl rejects the Aristotelian division of the elements into four kinds: earth, water, fire and air. Bodies are divided instead into simple principles (water, earths) and compounded principles (compounds, mixts, aggregates). However, as Kuhn pointed out (1962) the term ‘principle’, unlike ‘principium’ or ‘principe’, does not necessarily mean basic and indivisible. Stahl’s principles are in fact basic substances out of which compounds are made; they are like our elements in some sense, even though some are characterized more in terms of what they can do (active functions) rather than in terms of what they are (as substances). The simple principles as indicated above include water and earths; and there are three kinds of earths: the vitrifiable principle, the liquifiable principle, and the inflammable principle or phlogiston. As for the compounded principles, they include compounds (lead, sulphur, mercury, etc), mixts (gold, silver, copper, etc), and aggregates (fire, air, etc). A mixt is defined as a body which consists of simple principles, while a compound is defined as a body which may consist of mixts.
The properties of compounds are explained in terms of the principles they contain. Phosphorus and sulfur burn because they contain the inflammable principle, phlogiston. In general, it was assumed that substances which burn are rich in a 'principle', phlogiston, which is imparted to the air upon combustion. So, when for example wood, wax, oil or charcoal is burnt, phlogiston is emitted into the air. Similarly, when a metal is heated, phlogiston is emitted, and we obtain a calx of metal (Stahl, 1723; Partington, 1961; Leicester and Krickstein, 1952; see also Thagard, 1990, p.187).

A more developed analysis of Stahl's conceptual network would have to exhibit the various relations holding between the elements (cluster terms or concepts) of his cluster conceptual network, between the network and possibly other cluster terms or concepts, and finally reveal, despite its incongruity to us, its coherence and explanatory power (Kuhn, Lecture I, 1988).

An even more detailed analysis over time would have to trace the development of this conceptual network through the works of Priestley (Philosophical Transactions, 1772; Experiments and Observations on Different Kinds of Air, 1774-86/1970; Considerations on the Doctrine of Phlogiston, 1796/1929), Cavendish (Experiments on Air, 1766/1961), Scheele (Chemical Treatise on Air and Fire, 1777/1931), and
Kirwan (*An Essay on Phlogiston*, 1784/1968). For these chemists have, as "veteran phlogistonists" further refined it and increased its explanatory power and coherence. It would have also to show how it enabled proponents of phlogiston theory to explain phenomena of combustion, calcination and respiration. Thus, it would have to explain for example why they held and believed the following proposition:

\[
P(i) \text{ Metal } + \text{ air } \overset{\text{heat}}{\rightarrow} \text{ calx } + \text{ phlogisticated air} \\
\text{ie. rich in phlogiston}
\]

And subsequently, why they held and believed the following cluster-structure underlying (i), whereby calx and phlogiston are considered parts or components of metal:

\[
P(ii) \text{ Phlogiston Theory [PT]}:
\]

\[
\text{Metal} \quad \text{calx} \quad \text{phlogiston}
\]

In passing, it would also have to explain why respiration was subsequently viewed as having the effect of removing phlogiston from the body into the air; and subsequently, why it was believed that if air is saturated with phlogiston by combustion or breathing, further respiration becomes impossible (Conant, 1964, p.70; Thagard, 1990, p.188).

Since some proponents of phlogiston theory knew also that combustion ceases in an enclosed space, it would have to explain how they accounted for this phenomenon.
Priestley (1774–86/1970) for example assumed that air has a limited capacity for absorbing phlogiston; he conducted his famous experiment which consists in heating red calx of mercury on its own; he found that he could obtain the metal of mercury and a new kind of "air" ['eminently respirable air'], which he called *dephlogisticated air* (1970, vol.II, pp.161–2).

P(iii) Red Calx ←<heat>→ mercury + *dephlogisticated* of mercury air (ie. poor in phlogiston).

He explained that the calx of mercury has been turned into the metal of mercury by taking up phlogiston; since the phlogiston must have been taken from the air, the resultant air must be de-phlogisticated. Dephlogisticated air supports combustion (and respiration) better than ordinary air, but somehow this was expected, since the removal of phlogiston from the air leaves the air with a greater capacity for absorbing phlogiston.

Phlogiston theorists (such as Cavendish, Scheele) also became interested in the properties of a gas, which Cavendish called *inflammable air*, and which he obtained by pouring a strong acid (sulphuric acid) over a metal (1766/1961, p.19). They upheld therefore:

P(iv) Metal + acid ----> salt + *inflammable air*.

Interestingly, the phlogistonists believed that the term 'phlogiston' designated the substance emitted during combustion. But, as the historical record suggests, they
also believed (at least for a time) that inflammable air which they had isolated was phlogiston. Once they had made this identification, they went on to record the properties of inflammable air (which was in fact hydrogen) using the term "phlogiston" (See Priestley's letters [1782-4], in Schofield, 1966, esp. letters #93, 95, 98, 114, 115; see also Kitcher, 1978, pp.533-4, note 17). This point only goes to support the notion of "field of referential possibilities" associated with a term-type. It also confirms the conceptual complexities suggested by the model proposed in chapter 1 [B]: Table 3.

A fully developed account would have to explain all of the above, i.e., the various changes undergone by the phlogiston conceptual network (i.e., additions, deletions, and/or restructurations, redefinitions of relations). But most importantly, it would have to explain how, as these changes were taking place, Lavoisier's conceptual network was progressively emerging as well, and finally came to replace phlogiston theory by 1789. It would have to explain in details the reasons which prompted Lavoisier to reject progressively the basic claims and explanations of phlogiston theory [e.g., P(i)-P(iv) mentioned above], and the mechanisms and processes by which he has come to establish his own conceptual network, oxygen theory, which laid down the basis for the beliefs and claims of modern
chemistry [M(i)−M(iv) below]. In short, phlogiston theory maintained that the processes here of interest all involved the removal of phlogiston, whereas Lavoisier's theory held that these processes all involve the absorption of oxygen. Furthermore, contrary to the structure assumed by phlogiston theory [P(ii)], Lavoisier's theory considered metal and oxygen to be parts or components of calx or rather metal oxide. Thagard is right in pointing out this particular structural difference between the conceptual networks here in question (1990, p.195). Modern chemistry's claims can be summarized as follows:

M(i) Metal + air −<heat>−> metal oxide + air/
poor in oxygen

M(ii) Oxygen Theory [OT]:
Calx/Metal Oxide
\[ \text{[metal]} \uparrow \text{[oxygen]} \]

M(iii) Oxide of mercury −<heat>−> mercury + oxygen

M(iv) Metal + acid −−−> salt + hydrogen.4

In order to develop a cluster account of how we went progressively from [P(i)−P(iv)] to [M(i)−M(iv)], it might be helpful to consider, as the historical record suggests, Lavoisier's conceptual network at four stages of its development: (1) Early Views, 1772–4; (2) Developing Views, 1774–6; (3) Developed Views, 1777–9; and (4) Mature Views: Oxygen Theory, 1780–9. [See Guerlac, 1961; 1975, pp.76–107; see also Thagard, 1990, p.189]. To do so would enable us to
follow Lavoisier's developing conceptual network, from one still "trapped in" the phlogiston framework to some extent, and which entertained rather tentatively and vaguely a different hypothesis [namely, that 'air might be absorbed in calcination and combustion'], to one which offered a real and systematic alternative to phlogiston theory.

As a preliminary, it is worth pointing out that Lavoisier came to chemistry from geology, mineralogy, and botany. This seems to fit with the common depiction of a "revolutionary scientist" as somehow an "outsider" to the field of research in which he or she brings about radical changes. Thus, he must have learned somehow how to play the "established game", but according to slightly different "rules" (or "examplars") than most other scientists, who are well-entrenched in the field, familiar with its game and rules. Also, the scientific community that he finds in place and of which he is to become a distinguished member is a well-structured community with already a relatively sophisticated infrastructure of communication networks (i.e., Academies, journals, training-schools, labs, societies, etc).

**Lavoisier's Early Conceptions, 1772-74**

Operating under a strict maxim which he imposed upon himself, "never to advance but from what is known to what is unknown and to draw no conclusions which are not warranted
by experiments," Lavoisier began to be interested in combustion and calcination in late 1772 (August/Sept). He undertook a series of experiments which were in his own words the starting-point of his entire work. Looking back, he wrote: "In 1772, I had already conceived the whole system that I have since published on combustion and calcination" (Lavoisier, 1862, vol.IV, p.86; my translation).

According to Guerlac (1975, pp.76-80), the historical record strongly suggests that at least two experiments were particularly important in confirming Lavoisier's suspicion that 'air' or some constituent of air, played an important role in the processes of combustion and calcination (see also the pioneering work of Meldrum on Lavoisier's first experiments ['Lavoisier's Three Notes on Combustion: 1772" 1932, pp.15-30].

Lavoisier's interest in the chemistry of air was aroused around 1771 by other French chemists that he trusted and respected (Guerlac, 1975, pp.77-8). Thus he came to know of the work of British pneumatic chemists (Black, Priestley, Hales) and of their results. For example, that "air is found in almost all bodies in nature, in almost all minerals, although perhaps not in metals" (Guerlac, 1975, p.77). Early in 1772, Guyton de Morveau published a book which proved conclusively that the well-known gain in weight of lead and tin when they are calcinated is not a
peculiarity of those metals, but that all calcinable metals become heavier when transformed into a calx; moreover, that this increase has a definite upper limit characteristic of each metal. De Morveau sought to explain these results by some invoking some fanciful variation of the phlogistic network, but Lavoisier suspected already that fixation or absorption of 'air' might be the cause.5

In August of 1772, Lavoisier decided to conduct a series of experiments in order to determine the role that air might play in chemical reactions involving metals. The question before him was: Is air absorbed or released from a metal when exposed for example to the strong heat of a burning glass? By September 1772, Lavoisier carried out his famous experiment on the combustion of phosphorus and sulfur. He discovered that the products of combustion weigh more than the original substances. In an unusual and memorable move on November 1, 1772, he sent a sealed note ('pli cacheté') to the French Academy of Sciences, which was not to be opened until he so desired, and that was on May 5, 1773. In it, he stated:

Sulphur, in burning, far from losing weight, on the contrary gains it...It is the same with phosphorous; this increase of weight arises from a prodigious amount of air that fixed during the combustion and combines with the vapors...What is observed in the combustion of sulphur and phosphorus may well take place in the case of all substances that gain weight by combustion and calcination; and I am persuaded that the increase in weight of metallic calces is due to the same cause [fixation or absorption of air]...This discovery seems to me one of the most

On the one hand, Lavoisier believed then that:

L(i) Metal + air $\xrightarrow{\text{heat}}$ calx of metal/with fixed or absorbed air.

L(ii)
\[ \text{\underline{\text{calx}}} \]
\[ \text{metal} \quad \text{fixed/ absorbed air} \]

But soon after this experiment, Lavoisier carried out another experiment using a modification of Hales's pedestal apparatus (an instrument used at the time to measure the amount of air released or absorbed). He heated red lead in the presence of charcoal in a closed vessel and observed that a large amount of air was given off. And so, he also believed:

L(iii) Red (calx of) lead $\xrightarrow{\text{heat}}$ lead + 'air' /\charcoal\checkmark

In a note dated August 1772, Lavoisier wrote: "A good number of experiments appear to show that air enters into the composition of many minerals, even of metals, and in abundance" (Lavoisier, August/1772; Guerlac, 1961, appendix, p.215; my translation).

One of these other experiments consists in pouring acid on a calx of metal, or in placing such a calx in acid (contrast with Thagard’s point in this respect, 1990, p.189). Lavoisier observed that in such a case, effervescence or emission of air occurs (compare with P(iv)
above of phlogiston theory). In another of the 1772 notes, he writes: "Effervescence is nothing but an emission [degagement/emanation] of the air that was somehow contained in each of the bodies" (Lavoisier, August 1772; Guerlac, 1961, appendix, p.215; my translation). And so, Lavoisier must have believed as well:

L(iv) Calx of metal + acid ----> 'effervescing calx'/<'air'/>?

In summary, one could say that by 1774 Lavoisier held and believed L(i)–L(iv):

L(i) Metal + air --> calx of metal/with fixed or absorbed air.

L(ii) \( \text{calx} \)\( ^{'} \text{metal}^{'} \)\( ^{'} \text{fixed/}^{'} \text{absorbed air}^{'} \)

L(iii) Red (calx of) lead --> lead + 'air'/<charcoal>

L(iv) Calx of metal + acid --> 'effervescing calx'/<'air'/>?

A representation of the relevant parts of Lavoisier's conceptual network at this point must include the following cluster-terms, relations, and explanatory claims or hypotheses:

(a) There are three kinds of substances: metals/air/calxes.
(b) Metals gain weight when they become calxes.
(c) Heated calx (in presence of charcoal) gives off 'air'.
(d) Effervescing calx produces 'air'.
(e) Calxes might contain 'air'?
(f) [(e) might explain (b), (c), and (d) ]?

It must be clear that Lavoiser's conceptual network at this point was only entertaining rather hesitantly hypothesis (e)
as a possible explanation for some 'puzzling' experimental results. In order to account for (b), phlogiston theory posited a bizarre notion of "negative weight," in spite of the fact that phlogiston had not been and could not be isolated (Partington, 1957, p.88).

Lavoisier seems to have been fully aware of the revolutionary implications of his experiments. On February 20, 1773 he is reported to have entered in the first of his famous lab notebooks (registres de laboratoire) that he intended to carry out "a long series of experiments (on the air which is absorbed by substances e.g., metals during combustion, or emitted from them during various chemical reactions) destined, as he put it, to bring about a revolution in physics and chemistry" (Lavoisier, 1773; Guerlac, 1975, pp.80-1; Mckie, 1965, p.xix). And yet, at that time, he was still under the 'sway' of the phlogistic network, still groping for his own. In this respect, a bit of focused historical exegesis might be appropriate. In draft memoirs titled and dated respectively Essay on the Nature of Air, August 1772, and On a New Theory of the Calcination and Reduction of Metallic Substances...April 1773 (relatively recently found and published by Rene Fric in Archives Internationales d'Histoire des Sciences, 12, 1959/1960, pp.125-139) Lavoisier seems to be developing two ideas at the same time:
[1] (All bodies can exist in the three states depending on the amount)... of fire which a substance combines with;

and [2] that the heat and light appearing during combustion derives largely, if not exclusively, from the release of the matter of fire combined with air, not from an inflammable principle in the combustible (Lavoisier, 1772/3; Fric, 1959, p.137; italics added).

Clearly, it is the second idea which was to provide Lavoisier with an important argument in his later open and deliberate assault on the phlogistic network. However, one can also see that his hypothesis arises somehow out of a phlogistic framework. As Fric puts it:

We can see that Lavoisier is still in 1773 a 'phlogistinian', but we can also feel that he is beginning to have doubts about Stahl's theory and its applications, and to perceive other possibilities for explanations (1959, p.137; my translation).

To use Kuhn's terminology, for a moment, we could say that Lavoisier had a clear sense of "anomaly" and a personal feeling of "crisis." In the April 1773 memoir, he writes:

My experiments are not complete enough yet to dare enter the lists against the famous chemist [Stahl]. Yet, I believe I have said enough that one can sense that the present theory of the chemists is defective in many points, and that it is probable that the phenomena of fixed air when more thoroughly studied will lead this science to a time of almost complete revolution (Lavoisier, 1773/1959, p.77; italics added; my translation).

Thus, although Lavoisier had perceived anomalies in the phlogistic network which created a crisis, at least for him, quite some time passed before he was able (around 1783) to launch and sustain a systematic attack, and foment a larger crisis in the scientific community.
Lavoisier's Developing Conceptual Network, 1774-6

In the early months of 1774, Lavoisier repeats a number of experiments and sums up all the relevant results in a report, *Opuscules Physiques et Chimiques* (1774;1776 [English Translation]; 1970). In it, he elaborates on the results of those experiments having to do with combustion and calcination, which seem to support the existence of "some sort of air" in metallic calxes, which he calls "elastic fluid." At this point in time, though further developed, Lavoisier's conceptual network does not enable him however to decide what this "elastic fluid" is: for example, whether it is a constituent part of air (he speaks of "an elastic fluid of a particular kind which is mixed with air") or air itself 'in its entirety,' as he sometimes says [*L'air par lui-même en entier*] (Lavoisier, 1774/1970, p.340). In other words, is it L(i) or L(i')?

L(i) Metal + air →<heat>→ calx of metal/with sort of air elastic fluid

L(i') Metal + air →<heat>→ calx of metal/with air itself

Again, this episode serves to illustrate the notion of "field of referential possibilities." Interestingly enough, it is reported that some weeks before Lavoisier presented such results as L(i') to the Academy, he received Priestley in Paris and heard him talk about his discovery of a kind of air, i.e., "dephlogisticated air" (Partington, 1961; Guerlac, 1975, p.86). As pointed earlier, Priestley had isolated
this new kind of air in August 1774 in his celebrated experiment in which red calx of mercury is (unusually) reduced to its metallic state upon heating without the addition of charcoal. Subsequently, he believed as noted previously:

\[ P(iii) \text{ Red calx/mercury} \rightarrow \text{<heat>} \rightarrow \text{mercury + dephlogisticated air} \]

In November 1774, Lavoisier resumes his experimental work and sets out to confirm and extend Priestley's experiment. Guerlac reports that when preparing the air from the mercury calx, he referred to it in his lab notebook as 'the dephlogisticated air of M. Prisley [sic]' (1975, p.86). Yet, in April 1775, in a memoir read to the Academy, he states surprisingly that

the principle which combines with metals during their calcination, which increases their weight, and constitutes them in the state of calx, is neither one of the constituents parts of the air, nor a particular acid diffused in the atmosphere; (...) it is the air itself in its entirety (Lavoisier, 1774/1970, p.408; italics added; my translation).

Apparently, he had been confused by Priestley's otherwise valuable and highly significant communication. How should we account for Lavoisier's temporary confusion? He clearly seems to have forgotten about his own earlier view, namely, that air is a mixture, perhaps even a compound, and that the metals on calcination combined with one of its components.

In any case, by 1774-1775, Lavoisier's conceptual network has changed in some significant ways (new
experimental results, new concepts and new relations): phosphorus and sulfur gain weight in combustion, just as metals gain weight in calcination -- by absorbing "air" [?]. In the Opuscules, Lavoisier also remarked that phosphoric acid was "in part composed of air, or at least of an elastic substance contained in air" (Guerlac, 1975, pp.90-1). Though he has acquired some of the structure of his later network, he is still not firmly convinced of the superiority of his network over the phlogiston network (Priestley's), or for that matter, that it can provide a real alternative to phlogiston theory.

Lavoisier's Developed Conceptual Network, 1776-8

Two years after the Opuscules, Lavoisier returns to the subject of air. On April 20, 1776, he presents "one of his most brilliant memoirs" to the Academy (Guerlac, 1975, p. 91), tilted "On the Existence of Air in Nitric Acid." In it, he refers back to his experiments on phosphorus and sulphur, and his conclusion that air entered into the composition of the acids which result from the combustion of these substances. Subsequently, he was led to consider the nature of acids in general and to conclude that all acids were in great part made up of air; that this substance was common to all of them; and further, that they were differentiated by the addition of "principles" characteristic of each. And thus, by "applying experiment
to theory," he was able to confirm his earlier suspicion, namely, that it was "a constituent of air," "the purest portion of air" (Priestley's "dephlogisticated air") which entered without exception in the composition of acids, especially nitric acid. The experiment consisted in heating a known quantity of nitric acid with a weighed amount of mercury; the result was a white mercurial salt (mercuric nitrate) which later decomposed to form the red oxide (red calx of mercury) giving off 'nitric air' (air nitreux) or nitric oxide. Then Lavoisier further heated the red oxide, and as he already knew, it decomposed and produced metallic mercury and "the air better than common air" or "eminently respirable air." It should be noted that Lavoisier was careful throughout this process to separate the different gaseous fractions by collecting the air under bell jars over water. His practice was significant in that analysis was followed by synthesis: by putting together nitric oxide and "pure air" in the presence of some water, nitric acid was regenerated. And this confirmed that his experiment was right.

By March 21, 1777, Lavoisier seems to have completely recovered from his earlier temporary confusion and lack of confidence in his own conceptual network. He presented a memoir to the Academy "Mémoire sur la Combustion en General," which in Guerlac's view constitutes "the earliest
announcement of his theory of combustion and his first, albeit cautious, assault on the phlogiston theory" (1975, p.104). In that memoir, he showed that phosphorus burned in air combines with the "eminently respirable air" to form phosphoric acid. When the air is used up, the remaining air, which he called 'mophette atmospherique' (nitrogen) does not support combustion or sustain life (respiration). In the same memoir, he showed that sulphur takes up "eminently respirable air" in the formation of vitriolic acid.

Thus, it is clear that by that time Lavoisier has finally developed a real and convincing alternative to phlogiston theory. Scheele, Priestley and Lavoisier were able almost at the same time to isolate the constituent part of air here in question (Wolf, vol I, 1961, pp.368-9). Scheele called it "empyreal air." Priestley, in line with phlogiston theory, called it as we have seen, "dephlogisticated air" and explained combustion and respiration as processes by which phlogiston is removed. As for Lavoisier, he was now convinced that the element involved in calcination and combustion is a constituent part of air, 'the purest', 'the most salubrious part of air', the 'eminently respirable part of air.' Interestingly enough, the 1775 memoir (1774/1970, p.408) was revised and a final version published in 1778 with a conclusion which now reads
as follows:

the principle which unites with metals during their calcination, which increases their weight and constitutes them in the state of calxes, is nothing but the most salubrious and the most pure portion of air (Lavoisier, 1862, vol.II, p.123; my translation and italics added).

We could then say that Lavoisier holds and believes:

L(i) Metal + air $\rightarrow$ calx of metal + air without 'its most salubrious part'.

L(iii) Red calx of mercury $\rightarrow$ mercury + 'pure air' or 'most salubrious part of air'.

A full characterization of Lavoisier's conceptual network at this point in his career (1776-8) would have to include all the additions and deletions, or restructurations which have been made. Lavoisier offers such a characterization himself in his "Memoire sur la Combustion en General" (1862, pp.174; 225-233; see also Leicester and Klickstein, 1952; Knickerbocker, 1962). Briefly, one should note that Lavoisier distinguishes four kinds of airs: atmospheric air, fixed air, mophette (nitrogen), and eminently respirable air/pure air/most salubrious part of air. He criticizes the proponents of phlogiston theory for not having isolated 'phlogiston' (after all this time), and even point out that the existence of an alternative hypothesis may well undermine the very foundations of the phlogistic system. Commenting about his attack on the phlogistic system however, Lavoisier still seems to insist that his purpose was not "to substitute for it a rigorously
demonstrated theory, but only an hypothesis which seemed to be more probable, more in conformity with the laws of nature, and one which appeared to involve *less forced explanations and fewer contradictions*" (Lavoisier, Memoir 1777/ Knickerbocker, 1962, p.134; Thagard, 1990, p.192; italics added). But it was not long before he finally launched a systematic and definitive attack against phlogiston theory, and completely overthrew it.

**Lavoisier's Mature Conceptual Network: Oxygen Theory, 1780-9**

This is nowhere more apparent than in his "*Reflexions sur la Phlogistique*" (Lavoisier, 1783/1786 [?]; 1862, vol.II, pp.623-5), which McKie called 'one of the most notable documents in the history of science' (1935, p.220), and which constitutes a closely reasoned refutation not only of Stahl's doctrine but also of all latter-day phlogistonists. In it, he states at the outset that phlogiston is imaginary, that its existence in metals and in all combustibles is a baseless supposition, and that all the facts about combustion and calcination are better explained in a better way without it than with it; in short, that it is unnecessary and his theory is simpler. He writes

> I have **deduced** all these explanations from a simple principle, that pure air, vital air (*air vital*), is composed of a particular principle belonging to it and forming its base, and that I have named *principe oxygine*, combined with the matter of fire and heat. Once this principle is admitted, the main difficulties of chemistry seemed to dissipate and vanish, and all the phenomena
were explained with an astonishing simplicity. But if everything is explained in chemistry in a satisfactory manner without the aid of phlogiston, it is from that alone infinitely probable that this principle does not exist. (Lavoisier, 1783/1862, vol.II, p.623; italics in text; my underlining; my translation).

By now, we could say that Lavoisier holds and believes:

L(i) Metal + air $\xrightarrow{\text{heat}}$ metal oxide + air/poor in principi oxygen

L(ii) Oxygen theory [OT]: \[ \text{Calc}/\text{Metal oxide} \]
\[ \text{[metal]} \]
\[ \text{[oxygen]} \]

L(iii) Metal oxide $\xrightarrow{\text{heat}}$ metal + principi oxygen.

Since "eminently respirable air" is a constituent of so many acids, Lavoisier came to believe (wrongly, as it turned out) that it may play the role of "acidifiable principle." We have seen that when combined with certain substances, this air formed acids of various sorts; but we have also seen it forms calxes (or oxides) when combined with metals. And yet, Lavoisier was impressed mainly by the "acid-forming" characteristic of this air. That is why he now refers to the acidifying principle or base of "eminently respirable air" as the "principi oxygen" (oxygen principle), i.e., etymologically "that which begets acid". Lavoisier's use reflects his (wrong) belief that all acids contain oxygen (exceptions: hydrochloric acid; hydrocyanic acid). What he now calls "air vital" (presumably at the suggestion of Condorcet [Lavoisier, 1862, II, p.263; Guerlac, 1975, p.138 note 25]) is described in accord with his conceptual
network as a compound of "principe oxygine" and of the matter of fire and heat.6

In 1789 in the Traite Elementaire de Chime, Lavoisier uses the term "caloric" to refer to the matter of fire and heat. In this respect, it is significant to note that Lavoisier assumed that air must contain a principle of heat in order to explain why combustion produces heat — just like proponents of phlogiston theory assumed that substances which burn must contain an inflammable principle, phlogiston (Guerlac, 1975, p.105). But for Lavoisier, air, rather than the combustible substances, was the source of heat. In passing, it might be interesting to point out that the notion of "caloric heat" was accepted until well into the next century when the kinetic theory of heat was developed.

It should be possible to reconstruct Lavoisier's conceptual network in 1783, [where 'oxygine' was not an element] and then in 1787 and finally in 1789 [where 'oxygen' is an element along with other elements]. Such a reconstruction would show that by 1789, Lavoisier's conceptual network has developed to the point where it constitutes the very foundation for the modern framework, announced at the start of this analysis, i.e., [M(i)-M(iv)]. With such a reconstructed network in hand (Lavoisier, 1789), it should be possible to compare it with the conceptual network of phlogiston theory and with Lavoisier's in 1772.
What such comparisons would reveal is that significant conceptual shifts have occurred (over a period of time), resulting in the transformation of the entire field of chemistry, not just through additions of new cluster concepts and deletions of old ones, but also through restructurizations and extensive alterations of the relations presumed to be holding between these concepts. According to Thagard (1990, p.195), the relations most likely to be involved are kind-similarity/difference relations and part/whole relations [compare P(i)-P(iv) and L(i)-L(iv) and in particular compare and contrast P(ii) and L(ii)].

The following graphic summary might help in understanding the progressive, yet radical development in the conceptual change which led from phlogiston theory to oxygen theory:

PHLOGISTON THEORY

P(i) Metal + air --<heat>---> calx of metal +
phlogisticated air (ie. rich in phlogiston)

P(ii) Phlogiston Theory [PT]: Metal
        /
       /\  
      [calx] [phlogiston]

P(iii) Red Calx of Mercury --<heat>---> mercury +
dephlogisticated air (ie. poor in phlogiston)

P(iv) Metal + acid -----> salt + inflammable air
LAVOISIER, 1772-74

L(i) Metal + air --> calx of metal/with fixed or absorbed air

L(ii) \--Calx--\
   \   
   metal fixed/
   absorbed air

L(iii) Red (calx of) lead --> lead + 'air'/<charcoal>

L(iv) Calx of metal + acid --> effervescing calx/'air'/?

LAVOISIER, 1774-76

L(i) Metal + air --> calx of metal/
   with sort of air elastic fluid

L(i') Metal + air --> calx of metal/with air itself

LAVOISIER, 1777-79

L(i) Metal + air --> calx of metal + air
   without 'its most salubrious part'

L(iii) Red calx of mercury --> mercury + 'pure air'
   or 'most salubrious part of air'

LAVOISIER, 1780-9

L(i) Metal + air --> metal oxide + air/poor in
   principe oxygine
L(ii) Oxygen theory [OT]:

\[
\text{Oxygen Theory [OT]:} \quad \text{Metal Oxide} \\
\text{[metal]} \quad \text{[oxygen]}
\]

L(iii) Metal oxide \(\rightarrow\) metal + principo oxygine

L(iv) Calx of metal + acid \(\rightarrow\) effervescent calx/oxygine/

------------------------------------------------------------------------

MODERN THEORY

------------------------------------------------------------------------

M(i) Metal + air \(\rightarrow\) metal oxide + air/

M(ii) Oxygen Theory [OT]:

\[
\text{Oxygen Theory [OT]:} \quad \text{Metal Oxide} \\
\text{[metal]} \quad \text{oxygen]
\]

M(iii) Oxide of mercury \(\rightarrow\) mercury + oxygen

M(iv) Metal + acid \(\rightarrow\) salt + hydrogen.

------------------------------------------------------------------------

Note: As I suggested at the beginning of this study, a more complete summary would include the development of phlogiston theory through the works of its main proponents, parallel to the one given for Lavoisier's theory.

------------------------------------------------------------------------

A more detailed cluster analysis than the one that I have offered here would combine the kind of treatment given above and the one illustrated in Table 3, chapter 1 [B]. It would have to be more specific about the nature of the changes involved as one cluster conceptual network emerges and finally replaces another one. It would have to be more specific about the changes in the cluster of properties
associated with a given term (e.g., phlogiston, oxygen), in
the field of referential possibilities associated with it,
the relations holding between this cluster-term and other
cluster-terms in the relevant cluster network, about the
specific ways in which these changes take place (due to
additions of new cluster concepts, deletions of old ones, or
restructurations of both new and old concepts). Short of
such analysis, one can however conclude that the chemical
revolution here under consideration clearly involved
additions, deletions, replacements, but also, and most
importantly, restructurations or reorganizations of cluster
conceptual networks.
B. Developmental Problems

A more detailed and complete cluster analysis would have then to account with a greater degree of specificity for how these changes and transformations took place, i.e., for the mechanisms and processes (cognitive, social, historical and otherwise) by which these changes and transformations took place for a scientist (Lavoisier) and the community of scientists at large, in France, Britain, and elsewhere. In this next and last section, I would like to make a few fragmentary comments about these issues from the point of view of the new and revised cluster theory.

If we had considered only Stahl's [1723-30] (or even Priestley's [1766-1796]) conceptual network and Lavoisier's mature conceptual network (1789), it would have been tempting to say that a gestalt or radical change took place, that the networks are therefore incommensurable (cf: Shapere, 1984; Nersessian, 1987; Kuhn, 1988). But, as we seek to understand how Lavoisier finally arrive at his 1789 conceptual network, and trace its development through the various stages from 1772 to 1789 [as we trace also the development of its rival, the phlogistic network], we see a different picture emerging, one which suggests progressive small steps away from the phlogistic network, without denying that in the end there was indeed a radical change.

In the early stages of this development (and even later on), Lavoisier was familiar with the phlogistic network as
attested by his sometimes phlogistic explanations (Lavoisier, 1774/1970). He knew the phlogistic network and could apply it, even as he was articulating his conceptual network. However, by 1776-7, as we have seen, with greater confidence in his own conceptual network, he made the decisive shift. Holmes reports that even then Lavoisier caught himself in drafts and manuscripts using Priestley’s term “dephlogisticated air” which he then crossed (1985, p.107; see also Lavoisier, 1862, II, p.263). And yet, he developed his own conceptual network on the basis of his own discoveries and experimental results and those of other distinguished chemists of the time. [Priestley was unquestionably influential in this respect, however one wishes to look at it (see Partington, 1961; Guerlac, 1975, p.86)].

A more detailed cluster-based analysis must be offered of the cognitive mechanisms and processes used by Lavoisier (e.g., concept formation, forms of reasoning, special heuristics, experiments, generalizations, hypotheses) and which led to his articulation and development of a new conceptual network. However, in the meantime, it might be illuminating to note the following. For example, his original training in other fields than chemistry may have constituted an advantage —as suggested above. Furthermore, though he may have been a theorist whose merit lies in his
capacity to take over others' results and ideas, to combine and rearrange them, and by a true logical procedure to expound the true explanations of these results, he was also a great experimentalist. As we have seen, he carried out himself numerous experiments, many of which turned out to be "crucial," and repeated many other experiments conducted by others chemists (of his time, and previous times) which he often showed in a striking way to be inconclusive. Lavoisier's work was heavily experimental, systematically quantitative and made extensive use of instruments, such as the balance. Even though he was not the first to apply quantitative methods to chemistry, his use of such methods was deliberate and systematic.

Lavoisier often called himself a "physical chemist" and used physical methods which led him to assume, rather significantly for his purpose, that "when several causes or circumstances operate, it is necessary to remove them all but one, and investigate each separately" (Lavoisier, 1862, vol.II, p.656). Such methods led him also to assume the "law of conservation of matter" ('nothing is created, nothing is lost, everything is transformed'). Though he was not the first to formulate such a law (see Helmont, Boyle), he certainly considered it to be at the very foundation of chemistry. He writes in this respect:

Nothing is created in the operation either of art or of nature, and it can be taken as an axiom that in every
operation an equal quantity of matter exists both before and after the operation, that the quality and quantity of its principles remain the same and that only alterations and modifications occur. The whole art of making experiments in chemistry is founded on this principle: we must always suppose an exact equality or equation between the principles of the body one examines and those of that we extract (retire) by analysis. (Lavoisier, 1789/1790, p.140; 1862, I,p.101; Guerlac, 1975, p.119; my translation; italics added)

How did Lavoisier's conceptual network finally replace the phlogistic network? My analysis reveals that this was not just done by adding or by rejecting particular concepts and claims of the phlogistic network, but finally by challenging the entire structure and replacing it by 1776-7 with what was a well-developed alternative, based on a different hypothesis, which had by then an established explanatory power, coherence, problem-solving capacity and simplicity.

The picture here fits Thagard's characterization in three stages of the replacement of a conceptual network by another (1990, p.202). First, we have in the "foreground," phlogistic network --with a certain established connection to concepts, 'observations' and experimental techniques other than those implicated directly by this network. Second, we have the phlogistic network still in the "foreground," with its known and established connection, but we also have in the "background" the emerging though still hesitant structure of Lavoisier's new conceptual network, possibly with a similar connection as the previous network. Although
by 1776-7 Priestley and Lavoisier had very different and somewhat incompatible conceptual networks, they nevertheless shared and agreed on a vast number of concepts, observations and experimental techniques. **Third**, --once its explanatory power and coherence is established for Lavoisier, we have the full structure of his new conceptual network which "pops up" and moves more confidently into the foreground, "upstaging" as it were the phlogistic network and "pushing" it into the background without however eliminating it altogether --at least for a good while. For, whenever necessary, Lavoisier and his collaborators could think and talk within the phlogistic conceptual network and even translate claims from one to the other (see at a later point reference to debate between Lavoisier, his followers and Priestley and Kirwan).

Under this kind of analysis, we have seen how conceptual change is a long, difficult and protracted process. It took Lavoisier several years to articulate and develop his conceptual network (from 1772 to 1777, 1783, or 1789 [?] --depending on where exactly one wishes to place the decisive shift, or as Bachelard would put it, "la coupure"). And so, it should not be surprising if it took years as well for some phlogiston theorists (e.g, Kirwan) to pass from opposition to Lavoisier's conceptual network to its acceptance (Perrin, 1981, pp.40-63). Exceptionally,
Priestley maintained phlogiston theory until his death (1804) and never accepted Lavoisier's conceptual network.

In a letter to Benjamin Franklin around Feb. 1790, Lavoisier writes interestingly:

I believe, and a large number of scientists today agree with me, that the hypothesis accepted by Stahl and subsequently modified is false, that phlogiston in the sense that Stahl gave to this word does not exist... Young people whose heads are not filled with any system grasp it eagerly, whereas older chemists still reject it, and most of them find it more difficult to grasp and understand than those who have not yet studied any chemistry... French scientists are at present divided between the old doctrine and the new. I have on my side M. de Morveau, M. Berthollet, M. de Fourcroy, M. de Laplace, M. Monge, and in general the physicists of the Academy. The scientists of London and of England also very gradually abandon Stahl's doctrine, but the Germans still cling to it. (Lavoisier, Correspondences, 1790; Guerlac, 1975. p.110-2).

However, as Perrin has shown, by 1796, most scientists in France and Britain had adopted Lavoisier's oxygen theory and his conceptual network. And by the end of the 18th century virtually the entire scientific community had accepted Lavoisier's new conceptual system (1988, pp.105-124). In fact, in 1791 already, Lavoisier could say:

All young scientists adopt the theory and from that I conclude that the revolution in chemistry has come to pass (Lavoisier, 1791; Mckie, 1952, p.207; 1965, p.xxiv; italics added).

Chemists who were opposed to his conceptual network would nevertheless repeat some of his "crucial" experiments, and thereby come to acquire a familiarity and understanding of the new conceptual network. Others would attempt to use
the network and apply it in order to explain some phenomena, and thereby acquire an insight into its explanatory power and coherence. Most often however, and most importantly, it seems that it has been arguments pointing out the 'anomalies' (difficulties and contradictions) of phlogiston theory and seeking at the same time to establish the explanatory power, the coherence and simplicity of oxygen theory, which have led most scientists of the time to accept the latter (cf: Kuhn-Hempel debate over theory-choice; see also Thagard, 1990, p.204). Lavoisier and his collaborators offered such kinds of arguments on different occasions (in 1777, 1783, and in 1789), and furthermore, in response to objections by some of the last major proponents of phlogiston theory (Kirwan, Priestley).

In this last respect, one should note the publication in 1784 of a defense of phlogiston theory by Richard Kirwan titled *Essay on Phlogiston*. It was translated in French in 1788 (by Mme Lavoisier) with responses from Lavoisier and his collaborators (i.e., Guyton de Morveau, Berthollet, Fourcroy, Monge and Laplace) inserted between chapters (Kirwan, 2nd English ed, 1789/1968, see especially pp.56-7; 104-5; 115-7; 176-7; 281-3; 314-6). This project is rather significant because it shows Kirwan systematically objecting to the new oxygen theory, in defense of phlogiston theory; and it also shows Lavoisier and his friends systematically
undermining the attempts by Kirwan to defend phlogiston theory.

The interchanges between them seem free of major problems of communication and (in)commensurability — of meaning or otherwise. Rational debate and disagreement was apparently possible. Even rational acceptance of a new conceptual network was possible since Kirwan rejected phlogiston theory and adopted Lavoisier’s conceptual network by 1792. Through this argumentative exchange with the French chemists, he must have come to see and accept that their system had a relatively greater explanatory power and coherence than his. Also worth mentioning on this subject, is Cavendish’s *Experiments on Air* in which he suggests a clear procedure for translating (a ‘translation manual’) between his own language and that of Lavoisier (1761, pp. 35-8). Finally, when Priestley published his *Considerations on the Doctrine of Phlogiston* (1796), he was attempting as if for the last time to object systematically to the arguments of members of the "anti-phlogistic task force"; he was certainly aware of their arguments, for he dedicated his book to Lavoisier’s surviving collaborators who had inserted their critical notes in Kirwan’s work (Thagard, 1990, p.205). In his case however, rational debate has not resulted in a conceptual shift, and one may speculate as to why.
Some (more or less) plausible explanations might include any number of the following. [These explanations might by the way apply as well to any other scientist or group of scientists]:

(i) Reasons and arguments are never sufficient -- even though they may play a significant role; to every reason, there is always a counter-reason, and to every argument, there is always a counter-argument.

(ii) Lack of sufficient evidence in cases, and failure of the new conceptual network to answer all questions about itself. Priestley (1796) for example believed that the new conceptual network was based on only relatively few experiments (not enough), which could also be explained by the phlogistic framework.

(iii) While recognizing that the phlogiston conceptual network faces some difficulties and even makes contradictory claims [e.g., inability to isolate phlogiston, to specify its (negative) weight], Priestley may have justified his position by pointing out that the new conceptual network also faces internal problems (e.g., Lavoisier's 'acid theory' and 'caloric heat').

(iv) Priestley had one of the most elaborate and sophisticated (comprehension of) phlogiston theory; and as a result, he could appreciate and defend its explanatory power and coherence better than any other chemists of the time.

(v) Priestley may not have been exposed to, used or applied the new conceptual network enough to be able to appreciate its relatively greater explanatory power and coherence, greater problem-solving capacity, and/or simplicity.

(vi) His eminent status and position in the scientific community of the time may have hindered the process of 'conversion'. In this respect, there may be as some sociologists of science seem to suggest some regular connection between the status and position of a scientist in the community and his or her capacity to reject an established conceptual network in favor a new one: the higher and the more prestigious his status and position, the more difficult it is for a scientist to reject the established network.
(vii) Priestley was already in the second half of his life and career, and it may have been too "costly" or too difficult for him personally to radically put into question and reject the theory that he had spent most of his life and career building, defending and improving. Cf: Gaston Bachelard's explanation (1934; 1984) for why older scientists are less prone to accept a new and revolutionary conceptual network or theory. Compare with Kuhn's explanation (1970), taken from Planck, according to which a revolutionary conceptual network only triumphs because its opponents die. Recall also Lavoisier's comments on young chemists adopting his conceptual network much quicker than older chemists.

(viii) Priestley may have been driven by certain nationalistic dispositions even though I have found no such evidence (e.g., French/British rivalry; German opposition, as Lavoisier seems to suggest in his letter to Franklin).

(ix) Finally, by virtue of the metaphysical beliefs a scientist harbors (i.e., Newtonian materialism vs. Cartesian mechanism), it might be easier (as for Black) or more difficult (as for Priestley and Cavendish) to covert to a new conceptual network.

According to Thagard (1990, pp.205-6), explanations (iv) and (v) account for why Priestley never became an oxygen theorist. Though these explanations are plausible [(iv) more so than (v)], others seem equally, if not more, plausible. Furthermore, it seems rather unlikely that Priestley did not know much about the "dephlogisticated air/oxygen" conceptual network.

In view of the serious resistance and opposition that they faced (at least initially --there were few "converts" prior to 1785), Lavoisier and his collaborators formed an "anti-phlogistic task force" which set out to "convert" the scientific community to the "New Chemistry." In order to
achieve this goal, a number of instruments seem to have been carefully selected, some of which turned out to be more helpful than others. So in addition to their personal contacts, the weight of their authority, the evidence of their "crucial" experiments, and the explanatory force of their arguments, the members of the "anti-phlogistic task force" were helped in their task by: (a) the New Nomenclature (or Classification of the Elements) of 1787 (b) Lavoisier's Traité (1789), and (c) the establishment of a new journal (Annales de Chimie) devoted to the new chemistry. Let us see briefly why.

First, the nomenclature, which represented the combined efforts of some of the best chemists of France, served to unify the field of chemistry like never before, and to eliminate the terminological confusion which reigned until then (Crosland, 1962). For the first time in the history of chemistry, the name of a substance (e.g., copper) designated primarily its chemical composition. Since the composition of a substance is theory-dependent, the nomenclature was a reflection of the theory, and its acceptance implied acceptance of the theory. This was in fact made clear by Lavoisier himself, in his April, 18, 1787 presentation to the Academy:

A well-formed language adapted to the natural and successive order of ideas will bring in its train a necessary and immediate revolution in the method of
teaching and will not allow teachers of chemistry to deviate from the course of Nature; either they must reject the nomenclature or they must irresistibly follow the course marked out by it (Lavoisier, 1787; Mckie, 1952, p.120; italics added).

Since it was the only systematic nomenclature in chemistry, the new system offered a good instrument in the French Chemists' attempts to win "converts."

Second, the establishment in 1789 of the Annales de Chimie, provided an outlet for the proponents of the new conceptual network, particularly younger chemists and foreigners, who could not publish in theMemoires de l'Academie Parisienne. This was particularly important since de La Metherie, the editor of the major scientific journal of the time, Rozier's journal, the Observations sur La Physique, was a staunch opponent of the new conceptual network.

Third and finally, when Lavoisier published the Traite (1789), a new conceptual network and a research program were systematically presented, complete with a new and improved version of the nomenclature, and exhaustive descriptions of "crucial" experiments. In a letter to Benjamin Franklin, Lavoisier comments about his book as follows:

I tried (...) to reach the truth through the close linking up [enchainement] of facts, to dispense with speculation as much as I could, for it is a treacherous instrument which deceives us, in order to follow as much as possible the torch of observation and experiment (Lavoisier, 1789; Guerlac, 1975, p.111; italics added).
The Traité did not merely present a new theory of combustion, calcination and respiration, but, as I suggested, a new foundation for chemistry, laid out according to a unifying conceptual network. Its publication signified perhaps the first modern chemistry textbook.
Chapter 3: Residual Problems

1

Can a cluster theory and its complex semantical apparatus be extended and applied as successfully to other cases of scientific conceptual change, and in particular, to other revolutionary episodes in the history of science? E.g., Copernicus; Galileo; Newton; Darwin; Einstein? Can such a cluster theory offer an adequate account of the process of meaning-making in science in general? The short answer might go as follows. If I were convincing in making a good case in its favor throughout this project, then, I think, the burden must have been shifted somewhat somehow; it must now be shown why such a theory will not do. The next generation of papers "Against a Defense of Cluster Theories" must now target specifically the cluster theory articulated herein.

2

A cluster theory such as the one advocated here might need to be further refined and extended in light of contributions in the cognitive, historical and sociological sciences, for example to provide a better account of the cognitive, social and historical mechanisms and processes by which scientists develop new concepts and new conceptual networks, alter, modify and eventually replace old ones. Is this a disadvantage or an advantage? It can seen as either,
depending on what one wishes to emphasize, i.e., its possible dependence upon external contributions or its capacity to incorporate and integrate them. In accord with one of the main tenets of the New Philosophy of Science (#8), the emphasis will be placed on the latter.

3

In the final analysis, someone might object: "Your version of cluster theory is very similar to what other theories propose (e.g., Shapere's, Nersessian's, and Kuhn's [1988], or else in linguistics and the cognitive sciences, but in a different terminology i.e., semantic networks, semantic fields, lexical fields). And so, one may wonder whether you offer more than a better terminology." First, I hope to have shown that though similar in some respects to other theories, the cluster theory advocated here is different and new in many other respects (see Part IV). Second, one of its distinctive features and advantages is indeed its terminology, which is based essentially on the notion of clusters. Such a notion, one can argue, is linguistically plausible, ecologically sound and psychological real. By virtue of the structure that it calls forth in the encoding of information that we, as human beings, seek in our interactions with the world and our environment, it is arguably a cognitively efficient structure; it maximizes the ratio of utility and information
over cognitive costs [i.e., \( E = U \& I / CC \)]. On the other hand, it seems to respect rather than obliterate the sometimes "messy" or rather "fuzzy" aspects of this picture, i.e., interaction between human beings (with limited cognitive capacities) and the world. Furthermore, one has to acknowledge the plausibility of the following picture, based on the notion of cluster through and through:

```
Mind  <--------> Language  <--------> World
|          |          |
|          |          |
<cluster-concepts>  <cluster-terms>  <clusters>
(categories)
```

Someone else might object: "In fact, your cluster theory is in some sense so drastically different, in its assumptions and principles, from traditional versions of the same kind that one has to say 'it's only a cluster theory by name'." Even though I do not think so, my response would be: 'May be --Give it another name, if you wish, and give it the benefit of the doubt: consider its content and major claims'.

NOTES

* I wish to thank Richard Grady for suggesting such a pertinent and provocative title.

Introduction

1. I am assuming that this much is agreed upon by philosophers of science, almost thirty years after the publication of Kuhn's Structure of Scientific Revolutions [1962].

2. Despite differences and divergences in their respective reasons for holding such a view, they agree that science is a process and argue that it ought to be viewed and studied in this light.

3. Throughout this project, I distinguish between (i) meaning i.e., semantic characterization of the things referred to by a term, and (ii) 'meaning' -- i.e., meaning as use, whereby the 'meaning' of a term is "how it is used by members of a linguistic (or scientific) community." See Achinstein [1968] for the latter view; see Putnam [1962] for a view of 'meaning' as reference-fixer.

4. I.e., about "what there is," and about how we know "what there is" (cf: Quine, 1969).

5. One should keep track of the different uses of the notion of "cluster" in this project; first, as in the present case, we have the ordinary, non-technical sense of the term; second, we have the predominant use, namely the technical, theoretical and methodological, sense underlying the new cluster theory.

6. May be so, in the sense that it differs radically in its assumptions from previous (contextual or cluster) theories of meaning. It is however still firmly based on the notion of "cluster" as both a theoretical and methodological tool, while being flexibly open to further improvements and possible transfers from other disciplines.

7. Thus, in contrast with many of the previous theories which assumed one or the other position, often as a guarantee of their viability, the new cluster theory that I envisage does not make such an assumption. On the one hand, some held that given (i) incommensurability, we have (ii) relativism, and that this leads to (iii)
anti-realism (see Oddie, 1989); on the other hand, others held that by assuming (i) realism, we avoid (ii) relativism and (iii) incommensurability.

8. For a more detailed discussion of my view on this subject, see my paper "Quine's Indeterminacy: A Dramatic Source of Incommensurability?" Unpublished manuscript, Rice Univ., 1990.

9. These different approaches exemplify in other words different kinds of semantics: [I] Intensional semantics; [II] Referential or Extensional semantics; [III] Cognitive-Historical semantics, or Naturalized semantics, as some have called it. To the extent that one may object to [III] that there cannot be a third or other approach to semantics, other than [I] or [II], the same objection would apply to the new cluster theory. But as I will show, the new theory makes a third approach even more plausible.

10. This explains in part the title: clusters' last stand. Though my goal herein is to refute the entering wedge of the argument against cluster theories and to establish the conditions of possibility of a new cluster theory [With respect to certain scientific terms such as "copper," "metal," "oxygen," "acid," "electron," "tiger," "fish," etc., which designate physical objects/phenomena/stuffs], it should be noted in anticipation that it can in fact be developed in some detail to other kinds of terms, and even applied successfully to a particular case-study (e.g., the "chemical revolution"), as I will do in Part IV. [See McKinsey, "Against a Defence of Cluster Theories," [1980]; see also Putnam, 1973; 1975; Donnellan, 1977; Kripke, 1980 for specific arguments against cluster (or descriptivist) theories — as they are sometimes called].

11. Many of the objections against traditional contextual or cluster theories have been raised by proponents of reference theories, who have subsequently presented their own kind of theory as the only alternative. If I am able to show that the objections which motivated and justified references theories are met by the new cluster theory, this can only undermine (the reasons for) reference theories being the only alternatives.

12. But first, a caveat must be entered. To the extent that I will be engaging in "rational reconstructions" of various theories, I will inevitably make simplifications or generalizations which may not be fair to a particular
theory or other. Still, I think I have introduced sufficient nuances in each case to redress this possible source of flagrant inaccuracies.

13. I am assuming contrary to much of the literature that the views of Carnap [1956] and Kuhn [1962] are in fact much closer than has been assumed (see Jane English, 1978; Newton-Smith, 1981; Graham Oddie, 1989). Among proponents of CTM's, I could have also included Hempel [1950-2], Feyerabend [1962-65], and Quine [1953].

14. Several versions of cluster theory have been formulated over the years (e.g., Wittgenstein, 1953; Searle, 1958; Strawson, 1960; Gasking, 1960; Putnam, 1962; and Achinstein, 1968, etc.). Though my account seeks somehow to capture and encompass various versions, it will deal explicitly with the last three and focus on Achinstein's in particular, because the latter is relatively speaking more developed and detailed, and because, it represents in the words of some, "the best hope for classical intensional semantics"—despite being a failure which indicates the need for a different (i.e., reference) approach (Goosens, 1977, p.134), and according to others, "still a viable and plausible account" [?] (Suppe, 1977; 1989).

15. Perhaps Achinstein's? In fact, I think that these authors should be defending the kind of cluster theory that I will advocate. Though this theory shares with the cognitive-historical approaches a number of assumptions and claims, it is, as we shall see in Part IV, radically different in several other important respects.

16. See Nader Chokr, "Prescription vs Description in Philosophy of Science, or Methodology vs History: A Critical Assessment" [1986].

17. For an answer to this question, see endnote, Chap.3/Part IV.
Part I MEANING APPROACHES

Chapter 1 Motivation and Context

1. Perhaps it is because of this distinction that some have argued that the only kinds of semantics possible must be (i) intensional or (ii) extensional/referential.

2. This is just another way of expressing the asymmetry between sense and reference.

3. And even later by post-positivists, as attested by the next sentence.

4. This clearly was the view of Carnap [1947] and Hempel [1950-2].

5. See for example Losee, 1972.

6. As Grandy also pointed out to me in a conversation (Summer 1990).

7. Or the "whole of science"? — as Quine sometimes suggests. I am not sure at all what this means and how it would work.

8. Except perhaps the one based on "disciplinary matrix" and its components (see Part I, Chapter 2 and Part IV).

9. See Kuhn [1970], p.206 quoted below; see also "Objectivity, Value Judgement, and Theory-Choice" [1977]; Grandy [1983]; see also Part I, Chap. 2.

10. See Giere [1988], p.34; Doppelt [1978], who argues instead that the incommensurability of standards is more important.

11. That is why incommensurability may be regarded in some sense as a "philosopher's problem."

12. Assuming that in philosophy we are still as concerned with the kind of questions one addresses as we are with the answers one seeks —if not more.

Chapter 2 Kuhn/Postscript [1970]

1. This is essentially what I mean by "infectious semantic holism."
2. E.g., Ptolemaic/Copernican system; phlogiston/oxygen theory; Newtonian/Einsteinian mechanics.


4. This is also what Kuhn himself proposes.

5. But as we shall see in Part IV, such a conclusion is too simplistic in this particular case.

6. See for example Giere’s "Constructive Realism" [1986].

7. This looks like a consequence of Grandy's principle of humanity (1973) -- as we shall see in Part IV.

Chapter 3 Concept-Representations: Sets vs. Clusters

1. Some examples from ordinary language might be "bachelor," "grandfather," "square knot," "shot," etc.

2. Could we call "essentialist" the view according to which the meaning of a term is determined by a set or cluster of properties, only relative to a given determinant background context and a given stage of scientific knowledge and development? Does it make sense, as Leplin assumes [1988], to talk of "essentialism" when the meaning of a term is thus relativized. In order to claim that "essentialism is scientific," as Leplin does, one must either (i) assume a questionable conception of the nature of science and scientific inquiry or (ii) change the meaning of "essentialism." Concerning my characterization of "grades of essentialism," it might be interesting to determine if, and how it differs from Quine's -- in the context of his objections to the modalities [1953; 1969].

3. They had in mind is traditional contextual theories of meaning, or else traditional versions of cluster theory.

5. This is exactly the point which has been overlooked by most philosophical theories (contextual or cluster), and which has been stressed over and over again by Shapere, Nersessian, and Kuhn [1982-90] (see Part III).

6. As Nersessian quite rightly points out (1984; 1987).

7. I will confine my attention to those main objections which are specifically but not exclusively relevant to the scientific language.

8. In fact, even when the locus of meaning was situated in the community, individual linguistic competence presupposed epistemic access (cf: Achinstein, 1968). Thus, most (contextual or cluster) theories of meaning were individual-, rather than community-based. This may be due in part to Frege’s legacy. Recall in this respect that for Frege [1892, 1970], the meaning or sense of a term is “that which is grasped by someone who understands the term.”

Chapter 4 Traditional Cluster Theories

[A] Gasing

1. In line with mathematical logic and set theory however, one may well be interested in (one type of) classes and yet talk about them in terms of sets. In other words, translation from one to the other is possible. Hence, throughout this discussion, I will talk in terms of sets, only to make things easier, and I will consider that to each class (which can change membership over time) corresponds a sequence of expressions (extensionally invariant descriptions) which designate different sets at different times (1960, pp.1-2).

2. Since one can reasonably assume that the key terms used above are understood, except for focus, we need to define it before going any further. Thus, given a set of entities Sn defined by a relation R, where Rxy = “x is related to y by R” and Fy(Sn) = “y is a focus of Sn” [Read “E” as membership symbol] we say that

\[(y) \text{ Fy}(S_n) \text{ if and only if } (x) [((x \in S_n) \iff (R_{xy} v x = y)].\]

3. This is to say that given a set Sn (where the subscript n indicates the size of the set) defined by a certain relation R, at some time ti, where the following holds:

\[(I) (x)(y) \ (R_{xy} \iff R_{yx}) \quad \text{[Symmetry]}\]
(II) \[(x)(y)(z) \quad [(Rx \cdot Ry) \rightarrow Rxz] \quad [\text{Transitivity}]\]

(III) \[(y), Fy(Sn) \iff (x) [(x \in Sn) \leftrightarrow (Rx y x=y)] \quad [\text{Omni-focality}]\]

We can then characterize a cluster by stating:

(IV) \[Sn \text{ is a cluster, at some time } t_i \text{, iff given } R \text{ as defined above and one (any) member } (y) \text{ of the set as focus } Fy(Sn), \quad (x)(y) \quad [(x \in Sn) \cdot (y \in Sn)] \rightarrow [Rx y . Fy(Sn)].\]

4. It is worth noting that Gasking is interested in the notion of "cluster" not just as a basis for a theory of concepts and meaning, but also as a basis for a new metaphysics. See also in this respect: Strawson, 1960; Sprigge, 1970; Part IV Chap. 1 [A].

[B] Putnam

1. According to Putnam, Wittgenstein [1953] also introduced his notion of "family resemblance" for the same purpose.

2. While still part of an intensional approach to semantics, Putnam [1962] is already announcing the extensional or referential approach to semantics (Putnam [1973/1975]). His 1962 view is based on the distinction between (i) meaning = "semantic characterization of the things referred to by a term," and (ii) meaning = "as reference-fixer or extension-determiner" --even though ultimately it does not fix, but rather over-determines the reference, as we shall see. The second sense will be fully incorporated in the new cluster theory.

3. There are presumably some advantages to do semantics in terms of (i) reference rather than (ii) meaning (or intension), for between the two, reference is presumably clearer. However, the approach that Putnam [1962] takes here is somehow between these two approaches.

[C] Achinstein

1. As I pointed out earlier, it represents in the words of some, "the best hope for classical intensional semantics" --despite being in the end a failure indicating the need for a different (i.e., reference) approach
---extensional semantics (Goosens, 1977, p.134). In the words of others, it still represents "a viable and plausible account" [?] (Suppe, 1977; 1989).

2. One may quite legitimately wonder about Achinstein's interest in dictionary "definitions," as a philosopher. This question aside however, we should note that different dictionaries will emphasize different aspects of a given list --depending on the purposes and the audiences to which they are addressed.

3. The term "property" is used here in a very broad sense to refer to any single piece of information about an item or entity that might be included in the "definition" of the term used to designate this item or entity.

4. It is not always clear whether Achinstein's account is inclusive or exclusive of logical necessity and sufficiency. In one sense, as we shall see in a moment, it is clearly inclusive, but in another sense, it is also exclusive, since the notions here question are presumably problematic and inadequate.

5. Does it make sense, as Leplin assumes [1988], to talk of "essentialism" when the meaning of a term is thus relativized. In order to claim that "essentialism is scientific," as Leplin does, one must either (i) assume a questionable conception of the nature of science and scientific inquiry or (ii) change the meaning of "essentialism."

6. Note that by hypothetical cases and scenarios, I do not mean science-fiction cases and scenarios (cf: Putnam, 1975), but reasonable and plausible cases and scenarios which one could easily envisage on the basis of present scientific theory and the history of science.

7. Achinstein considers the question whether Bronsted's theory of acids, based upon "inner structure," reflects a change from the 18th century notions of acids, according to which the only known relevant properties were 'observational'. Achinstein says that aside from possible changes in centrality of relevance, the answer depends upon "whether, as the term was once used, and still is by many, having a common "inner structure" (though unknown) would have been considered semantically relevant for being an acid..." He allows that granting continued centrality, the observational properties were possibly (but not probably) always non-semantically relevant, and that the use changed only in being more specific.
8. He writes for example: "Such sentences, which attribute semantically relevant properties of X to X, could be construed as expressing analytic statements if they are construed as expressing statements appropriately defended solely by appeal to linguistic descriptions," i.e., to the use of the terms in those sentences. (1968, p.40-1).

9. This may have been one of the main motivation for Achinstein’s adoption of the notion of semantical relevance.

10. I see no difference ultimately between (i) the cluster theory which characterizes the ‘meaning’ of a term in terms of relevant properties associated with the term on linguistic grounds alone, and (ii) that which characterizes it in terms of semantically relevant properties (associated with the term on linguistic grounds).

11. A more complete linguistic description for an A-term will include considerations of (i) centrality among properties and clusters of properties; (ii) the question of unknown as well as known properties—which is somewhat problematic; and (iii) it will be generally speaking characterized by some indefiniteness, involving borderline cases (see Achinstein, 1968, pp.19-25; for details).

12. These include the following models: Model A: (a) There are minimal conjunctions of logically sufficient properties. (b) There is more than one such minimal set; accordingly we can speak of a disjunction of (possibly overlapping) sets, each disjunct being logically sufficient, the entire disjunction being logically necessary. (c) Any property in any such minimal set is semantically relevant. Model B1: (a) There is a set of properties, each member of which is assigned a numerical weight. (b) There is a rule to the effect that to be classifiable as X, an item must have properties the sum of whose numerical weights is at least some specifiable number. Model B2: (a) Properties are divided into general categories; each considered semantically relevant. (b) There is a rule that to be classifiable as X, an item must have properties the sum of whose numerical weights is at least N [specified such that an item may be required to have properties in more than one category]. Model B3: (a) Like B2, except that (b) is replaced with a new rule: to be classifiable as an X, an item must have properties the sum of whose numerical weights is approximately N [or N +/- e].

14. Achinstein attempts thus to avoid Quine’s strictures against meaning and analyticity [1953], and the indefensible charge of "essentialism of meaning." But as I show, he does not succeed in avoiding what I call "third grade essentialism" --even though he introduced a useful distinction [meaning (i.e., semantic characterization of the things referred to by a term) $\neq$ 'meaning' i.e., "meaning as use": the 'meaning' of a term is "how it is used by members of a linguistic (or scientific) community"].

15. Insofar as the semantically relevant properties associated with a term have to do with the semantical aspect of the use of the term, Achinstein seems to be assuming that the latter aspect is the most important and significant of the aspects of use (e.g., syntactical, pragmatic, semantical) [see note 16].

16. The three aspects of the use of a term "X" can be characterized briefly as follows: (1) Syntactical aspect: to recognize and construct grammatically or syntactically correct as well as deviant expressions containing the term "X." (2) Pragmatic aspect: to use the term "X" to do a thing or in the course of doing a thing, such as describing, reporting, explaining, questioning, requesting, etc. (3) Semantical aspect: to distinguish X's from other things, say Y's; to recognize prototypical or exemplary cases of X's; to apply the term "X" to appropriate items X's.

17. In the new cluster theory, it is recognized that there are different kinds and degrees of linguistic competence, that linguistic competence does not presuppose epistemic access, i.e., full and direct knowledge of the semantical aspect of use of the terms here in question.

18. In my reading of Achinstein, I find that he does make these assumptions. In general, I think that his account suffers from the fact it keeps too much of the positivists' claims, even though he clearly makes some departures from them, and attempts to incorporate some of the historicists' claims.

19. In some sense, this is still a problematic linguistic description, but it suffices at this point to show in a preliminary fashion some of the distinctive features of the new cluster theory (see Part IV for details).
Part II REFERENCES APPROACHES

Chapter 1: Scheffler

1. The list includes several other contemporary philosophers, who favor a reference approach (e.g., Field, Putnam, Kitcher, Newton-Smith).

2. Clearly Scheffler is alluding here to what I have called the "asymmetry" between sense and reference.

3. I am tempted to ask: do we really? Still, does this constitute an argument in favor of referential semantics, rather than intensional semantics?

4. As we shall see throughout Part II, it remains to be shown that to do so would not result in the same or worse problems than those encountered by meaning-based theories.

5. It is interesting to note that like most proponents of a reference approach, Scheffler appeals and assumes "stability or identity of reference," while others talk of "continuity of reference," or "commonality of reference" — as if reference could not itself change radically in some cases, as a result of a revolutionary episode in science. Even though one could not make the case that reference, like sense, changes inevitably as result of a radical theory change, one must envisage and account for (the various ways in which) reference changes, particularly as a result of a scientific revolution.

6. Variations on Scheffler's proposal can be found in articles by Michael Martin [1971, p.17-26] and Arthur Fine [1967, p.231-40]. Martin argues that the possibility of objective decisions between scientific theories will be assured not only when they have terms with identical extensions but in any situation where certain relations of inclusion obtain between the extensions of their respective terms. As for Fine, he maintains that statements of different theories will be sufficiently logically related for objective comparison whenever they contain expressions whose extensions coincide for at least some restricted range of objects. I shall not discuss these refinements as such, but the criticisms I make of Scheffler's arguments could easily be adapted to apply to these further suggestions.

7. As I will argue in Part IV, the relationship between sense and reference (or equivalents) is not strict and deterministic.

8. See for example case-study, Part IV, Chapter 2.
9. In the end, if the doctrine of conceptual relativism that Scheffler seeks to undermine in order to re-establish the objectivity of scientific debate, is that for any two theories competing in the same field and separated by a revolution, there are some terms or expressions in each theory whose referents are not specifiable in the other theory. And if one can argue that he cogently presented such a doctrine as a thesis about reference. One should note however that he fails to distinguish conceptual relativism from another position defended by Kuhn, namely, that in scientific revolution, the referents of some terms or expressions are changed. As a result, he chooses to undermine conceptual relativism by claiming stability of reference through revolutions. But, strictly speaking, as Kitcher notes (1978, p.521), referential change is neither necessary nor sufficient for conceptual relativism. In fact, in a trivial sense, conceptual relativism can occur without referential change if the theories or languages involved contain radically different terms or expressions. Perhaps, more importantly, even if some (or all) terms or expressions were to change in reference, this would not imply that there are some terms or expressions of theory or language whose referents cannot be specified in the other theory or language. Thus, Scheffler's approach to defend referential stability may commit him to a position which is on the one hand unnecessarily strong, and which, on the other hand, misses an important aspect of scientific change.

Chapter 2: Kripke/Putnam's CTR

1. It might be helpful to revisit the kind of example with which Kripke initially elaborated the causal theory of reference (hereafter CTR) in detail, namely proper names for persons, like "Aristotle," "Russell," "Quine," "Churchill," etc. He writes:

   Let us suppose that we do fix reference of a name by a description. Even if we do so, we do not make the name synonymous with the description, but instead we use the name rigidly to refer to the object so named, even in talking about counterfactual situations where the thing named would not satisfy the description in question. Now, this is what I think is in fact true for those cases of naming where the reference is fixed by description. But, in fact, I also think contrary to most recent theorists, that the reference of names is rarely or almost never fixed by means of description. And by this I mean what Searle says: 'It is not just a single description, but rather a cluster, or family of properties, that fixes the reference.' I mean that properties in this sense are not
used at all (Kripke, 1972, p.157; 1980; italics added).

As this passage suggests, the prevailing view (cf: Searle, 1958) before Kripke's proposal was that a proper name referred to the person it did by virtue of the user associating it with a set or cluster of descriptions; the person referred to would be whoever fitted or satisfied the set or cluster of associated descriptions. Kripke's view however is that we can refer to someone by a proper name even if we know no descriptive facts whatsoever about that person, and indeed even if we have a number of positively false descriptive beliefs about him. He holds such a view because he takes the reference of a proper name to be "rigidly" fixed as follows. Kripke supposes that there was an "original baptism" at or during which the person was picked out and dubbed with his or her name. Those present at the baptism, or "original dubbing ceremony" will then use the name in communication with others, and they in turn will pick it up and pass it on similarly to other speakers, and so on. So when any member of the community currently uses the name, his/her so doing will be the terminal point of a "causal chain" starting in the original baptism, and continuing, via other speakers' uses of the name, up to the present use. According to Kripke, it is simply this that makes the name in later speakers' uses refer to the same person, the person originally dubbed with it; the name's being there and used as it is constitutes the terminal point of a "causal chain" going back to that person. We should note, once again, that according to Kripke the later speaker need not know any descriptive facts at all about the person s/he is naming, need have no way of identifying the person in question. Indeed s/he may have only false beliefs about that person and consequently may be inclined to mis-identify him or her. Yet, in Kripke's view, that person would still be the person being referred to simply by virtue of standing at the beginning of the causal chain terminating in the later speaker's use of the name.

This account of proper names for persons seems to have some plausibility. According to some, it even fits our intuitions better than the traditional descriptive account (cluster or otherwise) and perhaps constitutes a definitive indictment of the latter. Such claims are usually backed up by the following sort of example. If someone thought that 'Quine' names the philosopher who wrote The Tractatus, and had no other thoughts about who that name stood for, we would not therefore be inclined to allow that when s/he said "Quine wrote The Tractatus" s/he said something true. But we should say so, if we thought that the person he referred to by 'Quine' was the person who best fitted or satisfied the sense or meaning (set or cluster of descriptions) s/he associated with that name. Note however that this does not
constitute, as some believe, a definitive indictment against any meaning or description-based account; it may at best be an objection against a particular version of such account, one which assumes that the sense or meaning, i.e., the set or cluster of properties or descriptions associated with a name, determines the reference, and is or must be known by the individual speaker using it to refer to a given person. For my present purposes, it shall suffice to raise some doubts about these assumptions and this kind of meaning or description-based account, by pointing out that it is possible to develop a meaning or description-based account in which the sense or meaning does not strictly determine the reference and the speakers' uses of a name are based rather on social and community standards, known to some groups of relevant experts and authorities, to whom individual speakers defer in their uses, knowingly or unknowingly. Such an account might be called a "social-descriptivist account" (see for example Marty Rice, Reference and Relativism, Dissertation, Ohio State University, 1987; see also Part IV, for how the new and revised cluster theory that I propose might incorporate the main insight of this kind of descriptivist account). 

2. Though Putnam is right in this respect, it must be stressed that any adequate theory should take into account and be constrained by "what goes on in the head" throughout the processes of meaning-making, i.e., the cognitive dimension.

3. The following passages from Putnam's work may help illustrate some of my characterizations and give a better idea of his view:

   If we put philosophical prejudices aside, then I believe that we know perfectly well that no operational definition does provide a necessary and sufficient condition for the application of any such word [i.e., natural kind words]. We may give an "operational definition", or a cluster of properties, or whatever, but the intention is never 'to make the name synonymous with the description'. Rather 'we use the name rigidly' to refer to whatever things share the nature that things satisfying the description normally possess (1975, p.238; italics in text).

For the fact that, in many contexts, we assign to the tokens of a name that I utter whatever referent we assign to the tokens of the same name uttered by the person from whom I acquired the name (so that the reference is transmitted from speaker to speaker, starting from the speakers who were present at the 'naming ceremony', even though no fixed description is transmitted) is simply a
special case of social cooperation in the determination of reference (1975, p.246; italics in text).

The extension of a term is not fixed by a concept that the individual speaker has in mind, and this is true both because extension is, in general, determined socially (...) and because extension is, in part, determined indexically. The extension of our terms depends upon the actual nature of the particular thing that serve as paradigms, and this actual nature is not, in general, fully known to the speaker (1975, p.245; italics in text).

4. Before turning to some discussion of these kinds of issues, it might be helpful to have a more perspicuous formulation of CTR. Let us assume that a theory of reference is simply a theory of what it is that decides for pairs \((t, r)\) whether or not \(t\) refers to \(r\). According to CTR,

\((1)\) \((1)\) has an "original referent" \(r^*\) \textit{(original sample)}.
\((2)\) \((a)\) \(r^*\) occasioned \(t\)'s "inaugural referential use," i.e.,
\((b)\) \(t\) was initially introduced as a name/term for \(r^*
\((3)\) \((a)\) Either \(r^*\) was experientially accessible (as for a natural kind like gold or copper) and \(t\) was introduced by ostension in a kind of "baptismal act" or "dubbing ceremony," or \((b)\) \(r^*\) was causally related to something else \(o\), which was experientially accessible (as in a cloud chamber track for "electron") and \(t\) was introduced as a name/term for the cause of \(o\). [NB: This clause is supposed to extend CTR to theoretical terms in science].
\((II)\) \((1)\) CTR furthermore assumes that a historical chain of linguistic behavior connects the present use of \(t\) to \(t\)'s inaugural referential use. \((2)\) The explanation of \(t\)'s present use cites a temporal and causal sequence of uses stretching back to the inaugural use such that at each post-inaugural stage of the sequence \(t\) is used with the intention to refer to its extant reference, whatever that be. That is to say, \((3)\) At each post-inaugural stage, the user of a term \(t\) need not be in a position to identify the referent independently of its use as characterized earlier.
\((III)\) Given \((1)\) and \((II)\), CTR decides that \(t\) refers to \(r\) if and only if \(r\) has some similar feature \(f\) which is an essential feature of \(r^*\).

5. In fact, as Shapere remarked, "it is clear that if this doctrine is correct, the continuity of scientific discussion is assured, at least in cases where a single term is used over a stretch of history, and the problem of incommensurability, of comparison of scientific theories within such continuous traditions is solved" (1989, p.423). Incidentally, it might interesting to note that much of Putnam's later work on the CTR, and its connection to the
problem of incommensurability, of theory comparison was based on his 1973 paper "Explanation and Reference," in which he stated:

[With a few possible exceptions (e.g., Feyerabend), realists have held that there are successive scientific theories about the same things: about heat, about electricity, about electrons, and so forth; and this involves treating such terms (...) as transtheoretical terms, as Dudley Shapere has called them (1969; 1984, chap. 17), i.e., as terms that have the same reference in different theories...The main technical contribution of this paper will be a sketch of a theory of meaning (rather of reference) which supports Shapere's insights (1973, p.200).

But, as we shall see in Part III, Shapere disagrees with Putnam's interpretation of the notion of transtheoretical terms.


7. Not unlike Scheffler in this respect, except that Putnam's theory is a causal theory.

8. A point which is often overlooked. See Shapere, 1984, 1989.

9. This only contributes to further undermine the viability of a causal reference theory.

10. Whatever the basic "constituents of matter" are, according to the best current scientific theory. See forthcoming point on this subject. Note also the "etc" i.e., the variability involved due to the open-ended nature of scientific inquiry, which a causal theory of reference overlooks.

11. Interestingly, as we shall see in Part IV, the same holds in the new cluster theory.

12. I don't see how a causal theory of reference can get around this.

13. Instead, it seems that a (causal) theory was built on the assumption of "commonality of reference" despite theoretical differences.

14. I think it is only fair to ask such a question of CTR.
15. And I might add in anticipation, correlated and assertible properties.

16. This is basically where a Putnam-like causal theory of reference fails, and where a new cluster theory becomes more viable.

17. See note 1 above. To return briefly to the point discussed in note 1, CTR's discounting of the descriptions or meanings associated with a term (or a proper name) may well fit with our intuitions (ref: example of Quine and the *Tractatus*), but this does not establish its viability for other kinds of terms; and this does not discount other theories from offering accounts which also fit our intuitions (viz., new cluster theory).

18. Though I have disagreed and objected to Putnam's CTR, I must however admit that I find it difficult to disagree with his conclusions. He writes interestingly enough:

   If there is a reason for (...) having gone so far astray with respect to [meaning] (...), it must be connected to the fact that the grotesquely mistaken views of language which are and always have been current reflect two specific and very central philosophical tendencies: (i) the tendency to treat cognition as a purely individual matter and (ii) the tendency to ignore the world, insofar as it consists of more than the individual's 'observations'. Ignoring the division of linguistic labor is ignoring the social dimension of cognition; ignoring what we have called the *indexicality* of most words is ignoring the contribution of the environment. Traditional philosophy of language, like much traditional philosophy, leaves out other people and the world; a better philosophy and a better science of language must encompass both (1975, p.271; italics in text; my underlining and addition in brackets).

Note in anticipation that the new cluster theory will not treat cognition a purely individual matter; it will incorporate a mentioned the principle of the division of linguistic and cognitive labor, and will therefore recognize the social or communal dimension of cognition and linguistic competence. And though it will also recognize the contribution of the environment, it does not uphold the *indexicality* of most words.

Chapter 3: Kitcher

1. See Introduction of this project for a similar point.
2. It is already a more nuanced view that I attributed to Kuhn.

3. I agree on the whole with Kitcher's claims, but I will disagree with their implementation.

4. He never really clarifies what he means by this. It is quite an assumption to make.

5. Kitcher makes I think a valid point.

6. In both these respects, Kitcher's account constitutes an improvement upon CTR.

7. It seems that Kitcher's account makes thus room for a cluster theory approach — at least partially.

8. It might be interesting to determine to what extent Kitcher's approach is similar/different, or complementary to Field's partial reference approach. But in the meantime, it shall suffice to note that it is hard if not impossible to find a convincing example from the history of science which would demand the application of Field's notion (see Chokr, "Field's Partial Reference" [1990]. Even Field's main example (i.e., "mass") does not seem to be a case which requires his approach (Earman and Fine, 1977, pp.535-8). And so, we'd better focus on Kitcher's approach: there may be indeed some advantages to understand claims about the development of scientific concepts as being in effect claims concerning the changes of reference potentials. (compare and contrast with my notion of "fields of referential possibilities"). It might also be interesting to determine to what extent Kitcher's approach (based on referential potentials) is similar/different to Newton-Smith's non-historical and non-holistic causal realist approach (1981, Chap. VII); see Chokr, "Newton-Smith's Causal Realism" [1990], for arguments why his approach fails to deliver an adequate account, based as it is on realism and the truth of the thesis of verisimilitude.

9. See part IV, chapter 1 [B].

10. That is, if one draws the full consequences of his view.

11. One of the objections raised by Shapere and Nersessian to Putnam's CTR is that it confines itself to science-fiction examples, rather dealing with actual cases from the history of science.

13. See Kuhn (1986), for a different conception of history; see also Part III, Chapter 3.

14. In order to further clarify his point and show what is involved in dealing with past scientific texts, Kuhn goes on to re-construct a text containing of some central aspects of phlogiston theory. And as he points out (1982, p.7), most of the terms in the sentences making up this text appear in both 18th and 20th century chemical texts, and they function in the same way in both. A few other terms in such texts, particularly "phlogistication," "dephlogistication," and related terms, can be replaced by sentences in which only the term "phlogiston" is not part of modern chemistry and its lexicon. But even after all replacements have been made, Kuhn claims quite plausibly, what we will find is that a few terms remain for which modern chemistry has no equivalent. In fact, as we know, some terms like "phlogiston" have completely disappeared from the language of modern chemistry. Others have a completely different sense. For example, as it has been repeatedly noted, a term like "principle" has lost all purely chemical significance. Still other terms, like "element," remain central to the chemical language, and somehow retain some of their older functions. But terms like "principle," previously learned together with them, have disappeared from modern texts of chemistry, and with them some previously constitutive generalizations have also disappeared. Consequently, the referents of these surviving terms as well as the criteria for identifying them are now radically different. In light of such an analysis, Kuhn concludes rather convincingly: Whether or not these terms from eighteenth century chemistry refer --terms like "phlogiston," "principle," and "element" --they are not eliminable from any text that purports to be a translation of a phlogistic original. At the very least they must serve as placeholders for the interrelated sets of properties which permit the identification of the putative referents of these interrelated terms. To be coherent a text that deploys the phlogiston theory must represent the stuff given off in combustion a chemical principle, the same one that renders the air unfit to breathe and that also, when abstracted from an appropriate material, leaves an acid residue. But if these terms are not eliminable, they seem also not to be replaceable individually by some set of modern words or phrases. And if that is the case, a point to be considered at once, then the constructed passage in which those terms appeared above cannot be a translation, at least not in the sense of that term standard in recent philosophy (1982, p.8).
In general, I think that Kuhn's point in this respect is well taken. If many, perhaps most, terms in older scientific languages and theories are identical in form and function with terms in a modern language or theory, others are new and must be learned or relearned together, as a components of a Cluster of correlated terms. These are the untranslatable terms for which the historian or philosopher has to discover or invent meanings in order to render intelligible the texts in which they appear. But interpretation (or language-learning, as Kuhn later suggests [1980]) and not translation, is the process by which the meaning or use of those terms is discovered or invented. Translation does not seem to arise. Thus, according to Kuhn, the task or activity of the historian or philosopher seeking to understand past scientific texts is perhaps better seen as falling under "hermeneutics," the "science of interpretation" rather than as "translation" as Kitcher and many other commentators have presumed (see Part IV, Chapters 1 and 2).
Part III: RECENT DEVELOPMENTS IN PHILOSOPHY OF SCIENCE

Introduction

1. In effect, they hold the main contentions of "the new philosophy of science" discussed in the Introduction.

2. Shapere is not very clear on this subject. However, to the extent that it is possible to flesh out the insight, we obtain, I argue, the new cluster theory that I present in Part IV.

3. A point to which I shall return in Part IV, Chapter 1B/1C.

4. This is the gist of his remarks in a 1990 paper, "Dubbing and Redubbing: The Vulnerability of Rigid Designation."

5. And hermeneutics, I must add. This may be explained in light of the institutional juxtaposition, geographically speaking, of the linguistics and philosophy departments at MIT, where Kuhn is presently teaching.

Chapter 1: Shapere

1. Clearly, Putnam was motivated by other considerations as well, but the problem of incommensurability seems to have been a major factor.

2. In a sense, the new cluster theory that I envisage in Part IV is also faced with a similar question.

3. Any adequate contextual or cluster theory of meaning must address this question, among others.

4. One may already object to Shapere that, like Putnam, he is interested in establishing continuity and commonality of reference, rather than allowing reference to change and vary, just as meaning changes and varies, across different theoretical contexts, and developing an apparatus to track and map out the changes (if any) in both meaning and reference. For more on this objection, see Part IV.

5. It is one thing to say that certain relations hold, it is another to say something about the nature of the relations. The same point can be made about Wittgenstein's notion of family resemblances.
6. An important point which will be incorporated in the new cluster theory, see Part IV.

7. This constitutes, as we shall see, the limitation of Shapere's model.

8. Generally speaking; though we are dealing here with the referential sense of "about," could we say, as Putnam claims, without further ado that Stoney, Thomson, Bohr, and later developments of quantum theory "referred all along to the same thing"?

9. That is, if it is one at all.

10. In my view, it is because, like Nersessian and Kuhn, he has in mind the traditional construal of cluster theory, and because he assumes that the objections against such a theory are well established and well-known.

11. It might be illuminating to contrast Shapere's characterization with that of Newton-Smith (1981, pp.159-161).

12. Interestingly, what seems to emerge from this passage is the idea that a concept is best represented by a cluster of properties, including the chain-of-reasoning connections linking the successive phases of development of the concept.

13. Along a line similar to Shapere's (1982), Nersessian writes: "Most responses to these problems have centered on discussion of the alleged nature and necessities of language per se; the presumption being that the results of such analysis can simply be transferred to the scientific case. Thus, actual linguistic practices in science, and in particular the processes of concept formation and change have gone largely unexamined (1987, p.161).

14. This point cannot be stressed enough, but Shapere does not show explicitly how.

15. See Part IV, Chapter 2 [case-study]; for how the new cluster theory explain this possibility with respect to the phlogiston/oxygen debate.

16. In effect, if this possibility is explained, then other cases of scientific change might be considered as particular, and less radical, cases of conceptual change.

17. But it needs to be developed and fleshed out, as I envisage in Part IV, Chapter 1 [B].
18. It is not so much whether it is restricted or enlarged which matters, but how its role is defined, i.e., strictly determinant or simply constraining, influencing, and shaping. The "solution" however is, to put it imply, to enlarge it, to make it much more "hodge-podge," and to define its role in the latter fashion, a shaping, influencing, and constraining.

19. Shapere's account may benefit from incorporating something like my principle of assertibility --relative to the current stage of scientific knowledge, or else relative to a particular background knowledge in a given domain, and a given community; see Part IV, Chapter 1 [A].

20. To the point of sometimes "flirting" with an extreme form of "historicism," as some commentators have pointed out (e.g., Nickles, 1985;1986).

21. Shapere, like Giere (1985), is in favor of a truly and genuinely "naturalized philosophy of science." There is no first philosophy or methodology from which to begin, nor is there a transcendental vantage-point. His slogan above is reminiscent of the "Neurath's boat" metaphor often used by Quine to characterize the enterprise of knowledge: fixing and building the boat while afloat.

22. These characterizations bring to mind the well-known criteria defended by the pragmatists (Peirce, James, Dewey).

23. Internalization and/or bootstrapping? Most likely, both processes are involved in some form or other. Shapere must however be given credit for having underlined and somewhat clarified the process of internalization.

24. Quine has also a similar tendency (1953, p.42), when he writes: "The unit of empirical significance is the whole of science."

25. By the same token, Shapere clears up the way for the new cluster theory that I will articulate.

26. The viability of Shapere's whole account, like Nersesian's for that matter, hinges on this question.

27. Shapere's point is well-taken. I will make a proposal on this matter in Part IV.

28. The "historicism" of Shapere's conception comes out clearly in this passage.
29. The new cluster theory makes room by the way for such a sense of possibility.

30. Rather than seeking desperately nice and abstract theories with crisp and clear principles, mechanisms, and structures which disregard or obliterate any actual messiness or fuzziness, we may have to break away from such well-entrenched philosophical tendencies and to do with theories which instead respect, and attempt to account for, the actual messiness or fuzziness found.

Chapter 2: Nersessian

1. She considers (i) logical positivism (ii) Kuhn and Feyerabend and (iii) Kripke/ Putnam's CTR; see for details Nersessian, 1984; Chapter 1 and 2; 1987, pp.161-4.

2. She writes interestingly enough: "In neglecting to add the dimension of 'discovery' to accounts of 'meaning' and 'meaning-change' in science, the prohibitions of logical positivism continue to influence contemporary philosophy" (1987, p.161).

3. The interest and concern with images, analogies, metaphors, and models in science is longstanding, but Nersessian proposes presumably a fresh look at their role and place in scientific concept formation and development.

4. As we shall see, Nersessian uses Gentner's structure-mapping theory (1983) for an analysis of the use of analogies in the formation of Faraday's and Maxwell's respective concept of "field." See Part IV for more on the possible connections and transfers between cognitive science (psychology) and philosophy of science.

5. In other words, though they have articulated the "historical turn," they have not presumably drawn the right consequences of this new stance. See Part I, Chapter 3.


7. Concerning the debate about a normative vs. descriptive approach in philosophy of science, see also Chokr, 1986, 1988.

8. An important point on behalf of the new cluster theory, as we shall see.
9. This is also by the way a problem of great concern from a historical point of view. For, as Kuhn stressed throughout his career, one of the historians’ primary concerns is to avoid the “Whiggish fallacy”, which I call the “retrospective fallacy,” i.e., attributing modern or present-day notions to past scientists. Concepts used in past scientific theories seem sometimes similar but they may differ radically from present-day concepts. Yet, according to Nersessian, historians do not seem to have a meta-theoretical account of the ‘meaning’ of a concept which would enable them to avoid this fallacy. In her analysis (1985) of the controversy over when Faraday had his ‘field’ concept, she shows that even the best historians are not immune to this fallacy, because they don’t have a clear articulation of what it is ‘to have’ a concept. And so, it is important to keep in mind that when we ask at what point did a given scientist ‘have’ a given concept, we assume not only that we have ‘criteria’ for the concept in question, but also that we have criteria for the form of representation of the ‘meaning’ of a scientific concept in general (see Nersessian, 1987, p.166).

10. As I will argue in Part IV, this list can and must be extended.

11. Again, see Part IV for further discussion.

12. Nersessian seems to be putting to use the notion of "schema" commonly used by cognitive scientists. Interestingly, Rumelhart, for example, characterizes a schema as follows:

A schema is a data structure for representing the generic concepts stored in memory. There are schemata representing our knowledge about all concepts: those underlying objects, situations, events, sequences of events, actions, and sequences of actions. A schema contains, as part of its specification, the network of interrelations that is believed to normally hold among the constituents of the concept in question....A schema theory embodies a prototype theory of meaning. That is, inasmuch as a schema underlying a concept (...) corresponds to the meaning of that concept, meanings are encoded in term of the typical or normal situations or events that instantiate the concept (1980, p.34).

13. Though the new cluster theory that I envisage is essentially an exemplification of the probabilistic view, it can in some cases accommodate the prototype view as well, or else its main insight, namely that prototypes play an important role. The only way to decide between (1) the probabilistic view and (2) the prototype view is to
consider cases of conceptual change. If (1) is the correct view of concept-representation, then one might find cases in which the prototype changed over time. In such cases, a different distribution of members satisfying the probabilistic criterion, i.e., the weighted properties criterion, a different member will be most (proto)typical. But if instead the prototype view is the correct one, i.e., if the prototype is fundamental and the properties weights derived, then we should find cases where the prototype remains the same, while the relevant properties or their weights change.

14. Some philosophers, very little impressed by Nersessian's proposal, have argued that not only it is not sufficient, but it exacerbates the problem (cf: Leplin, 1988). We shall consider in a moment his arguments.

15. See Nersessian, 1987, pp.165; for a summary of the different 'field' conceptions in the historical period stretching from Faraday to Einstein.


17. A quick survey of the literature would testify to that.

18. The new cluster theory that I will defend in Part IV incorporates these contributions.

19. To this extent, it will play an important role within the new cluster theory.

Chapter 3: Kuhn Revisited

1. See note 5 of the Introduction, Part III. Even though I will not be referring directly to Kuhn's Shearman Lectures (1988), I must say that they have significantly influenced my understanding of his latest work.

2. Clearly, Kuhn is interested in re-drawing the significance (theoretical and methodological) of the "historicist turn."

3. "Ethnographic" or "hermeneutic" for that matter.

4. This is also what the new cluster theory recommends.

5. Kuhn recognizes in passing that some categories in the lexicon are innate, i.e., genetically determined and shared by all human creatures, while others are species-universal
by virtue of shared aspects of the natural environment. And still other lexical categories have evolved in response to the developing needs of particular communities which may vary with both time and place, with the relevant community's environment, and with the way members of the community interact with that environment.

6. This reasoning could well be applied, as Kuhn in fact does (1970;1987), to the classification of celestial bodies in (i) Ptolemaic astronomy and in (ii) Copernican astronomy:

(i) The Sun, the Moon, and Mars are grouped into the same set.

\[ \text{Similarity} \langle \text{Sun, Mars, Moon} \rangle. \]
\[ \text{Dissimilarity} \langle \text{Mars, Earth} \rangle. \]

(ii) The Sun, the Moon, and Mars are classified in three different categories.

\[ \text{Dissimilarity} \langle \text{Sun, Mars} \rangle. \]
\[ \text{Similarity} \langle \text{Mars, Earth} \rangle. \]
\[ \text{New Category} \langle \text{Moon} \rangle \]

7. See note 6 above.

8. See Part IV, proposition [B] (b) of the new cluster theory.

9. I.e., how much commonality is required to explain the success of the historian in reconstructing a past belief-system, or that of another community?

10. In Kuhn's view (1990), lexical acquisition requires only that, for any pair of languages, overlap can be found in enough places to get the learning process off the ground. Nothing like a universal set of shared observational primitives is needed, and it is fruitless to look for one. The criticism of Carnap's program is obvious.

11. See discussion of Kitcher (1978) in Part II, Chapter 4; see also Part IV.

12. In fact, implicitly some form of such principle is applied by Kuhn throughout.

13. But only "local."

14. Alternatively, it can be formulated as the problem of determining whether the truth or falsity of a belief about the world depend upon the lexicon of the community within which that belief is held? In what way?
15. While Kuhn denies this charge, he recognizes that his view poses a threat to realism. He claims however that Putnam's "internal realism" bears instead significant parallels to his own position (1990).

16. Can talk of the other worlds of other communities be taken as anything but the wildest metaphors? -Kuhn asks (1989).

17. These expressions often led to asking whether Kuhn meant to say that there were *witches* in the 17th century, the *phlogiston* existed in the world of 18th century chemistry, or that *ether* existed in the early-20th century physics. Though Kuhn has sometimes answered such questions with a 'yes', it was "always in a most equivocal and embarrassed tone." Now he says in effect, he should have rejected these questions as ill-formed.

18. Only to recapture it from the point of view of the scientist(s) involved.
Part IV  NEW CLUSTER THEORY

Chapter 1:  Toward a Theory of the Process of Meaning-Making

Motivation/Context

1.  Recall that the main objections against traditional contextual or cluster theories of meaning are:
   (i)  Inadequate Concept- or Meaning-Representation
   (ii) Locus of Meaning: Individual vs. Community
        Linguistic Competence/Epistemic Access
   (iii) Infectious Semantic Holism/Essentialism of Meaning
        Paradoxes of Incommensurability.

2.  Contrary to what many philosophers believed (e.g., Putnam, Kripke, Shapere, Nersessian, Kuhn, etc).

3.  As I pointed earlier, if I am able to show that the objections, which motivated and justified reference theories as the only other alternative, are met by the new cluster theory, this can only serve to "rob away" and undermine the reasons for reference theories.

4.  In a sense, one could argue that I am trying to develop an alternative to Frege's view, but along the descriptivist tradition that he inaugurated. Instead of an individual-based descriptivist theory, I am proposing a community-based descriptivist theory using a cognitive and historical approach. Furthermore, instead of sense (or meaning) and reference, and the particular determining-relation holding between them, we have, as we shall see, a Cluster of properties and its associated field of referential possibilities [FRP]; and a new relation (of over- or under-determination) holds between a given cluster of properties and a referent.

5.  This is by the way an objection often raised by proponents of a reference approach.

6.  Originally, Putnam [1975] formulated the hypothesis of the division of linguistic labor. In a recent paper however, Kitcher talks of "The Division of Cognitive Labor" [1990]. Putting the two ideas together, I obtain the principle under consideration. More on this in forthcoming discussions.

7.  Also, as I pointed out in Part I, Chapter 4 [C], I don't see a difference ultimately between (i) a cluster theory which characterizes the 'meaning' of a term in terms of relevant properties --associated with the term on
linguistic grounds alone, and (ii) that which characterizes it in terms of semantically relevant properties --associated with the term on linguistic grounds alone.

8. My primary concern in this project will be with terms which designate physical objects, phenomena or stuffs, such as "copper," "oxygen," "metal," "acid," "electron," "atom," "tiger," "fish," "insect," etc. Ultimately however, I would want to claim that the new cluster theory can be extended, provided some modifications and adjustments are made, to other kinds of ordinary or scientific terms. In any case, even if the new cluster theory applied plausibly only to the terms here in question, its scope would be considerable and the theory would be worth contending with, at least as a partial account of (scientific) language, of the process of meaning-making --given that a good deal and an important part of (natural and scientific) language is made up of such terms.

9. More on this aspect of the new cluster theory is forthcoming.

10. This is only a preliminary and tentative representation, deliberately kept simple for better contrast with Traditional Contextual Theories of Meaning. But the new cluster theory is much more complex as we shall see in a moment.

11. Admittedly, we may have to clarify first what we mean by "reasons" and then the different kinds of possible reasons. But where should we turn for this purpose, if not the "context of science," where "we have learned how to learn"? and to give reasons of various kinds? (see Shapere, 1984;1989). Recall that the business of discovery is not only open-ended, but backbiting as well: our conceptions which serve this purpose are molded by it.

12. Because they overlook this, term-to-term translation methods, even of the context-sensitive kind as in Kitcher [1978], fail to deal satisfactorily with incommensurability problems (see Kuhn, 1982;1988 on this very point; see Part II; Part III, Chap. 3).

13. Here is graphically an example of the role of a multivariate (non-monolithic, non-determinant) background context:

```
------------> BC1 -----> Ti
|
BC =/= =/= =/= 
|
------------> BC2 -----> Tj
```
Though terms (or theories) $T_i$ and $T_j$ are different because they are shaped, influenced or constrained by different background contexts, $BC_1$ and $BC_2$, they (may) still have much in common by virtue of depending both upon a larger background $BC$, which encompasses $BC_1$ and $BC_2$ and much more.

14. Let us recall however that, unlike a reference theory which assumes "rigidity or fixed reference" (i.e., essentialism of reference), a cluster theory wants to account for the fact that both the cluster of properties and the reference may change over time. In this sense, it assumes neither an essentialism of meaning nor an essentialism of reference (see Shapere, 1982; Kuhn, 1986). Along with many others, here is one feature of the new cluster theory which may prompt one to ask if this theory is not a cluster theory only by name?" Possibly. But it is certainly a revised and reformulated version of cluster theory.

15. It seems to me that any adequate theory of the process of meaning-making must be "naturalized" in that it must reflect the minimal aspect of linguistic competence, the defeasibility of the core information or conception involved, and the communal dimension of language use. To the extent that the new cluster theory satisfies these requirements, it constitutes a proposal for a "naturalized cognitive semantics." Compare and contrast with Grandy, "Semantics Naturalized," unpublished, 1986, pp.1-11.

16. The historical record supports in fact these claims, as I found out in my investigations for the purpose of a case-study on the chemical revolution. See in particular the debates and discussions between Lavoisier/his collaborators and Kirwan/ Priestley, two prominent phlogistonists [Kirwan, 1784; Priestley, 1796]. Contrary to what has often been assumed, there were no breakdowns in communication, no insuperable kind of incommensurability. See Part IV, Chapter 2 [case-study].

17. Davidson has argued [1973;1984] that it is possible to justify transcendentally a principle of charity which ordains us to endeavor to maximize the ascription of true beliefs in the interpretation of the discourse of others. We should recognize however that we do not always can, have or want to be charitable in Davidson's sense. This is a point that many philosophers of science have neglected or overlooked. For, whatever plausibility the principle of charity has or may have as a sort of a priori constraint on
the interpretation of ordinary discourse of others, is not transferable to the interpretation of theoretical discourse. The simple reason is that we well understand how easy it is to have a theory which turns out to be totally incorrect: the history of science is filled with such examples. While it may be hard to see how a group of people could cope with the everyday world in the face of massively mistaken beliefs (this is what gives the principle of charity its plausibility), it is easy to see how a group can utterly mistaken at the theoretical level.

18. As we seen (in Part II, chapter 3), Kitcher claims that historians of science have in fact always tacitly assumed or used a principle of humanity. It is also worth noting that Kathryn Pyne Parsons advocates a similar principle and attempts to show how such a principle can be used in studying the history of science (1975, pp.367-96; 378-388 on the last point). Interestingly, what is known today as "hermeneutics" also makes use of such a principle, and possibly other similar ones (Kuhn, 1989). Finally, see A. Goldman's recent and unpublished "Interpretation Psychologized" [1990, pp.1-52] for another possible use of the principle here in question.

Representational Problems

19. See Nersessian, 1987, pp.166ff;169ff. See also forthcoming discussion.

20. In passing, it is worth noting that while Kuhn emphasized radical and discontinuous scientific change, Shapere/Nersessian have emphasized continuous scientific change.

21. After Kitcher [1978, p.540], we shall hold that when the FRP of a term is determined by one entity or type of entity, it is "homogenous." But when the FRP is determined by two or more entities or types of entities, it is "heterogeneous," or else, the term is "theory-laden."

22. For the case of "copper," it is worth recalling in this respect what the history of chemistry reveals: before the 17th/18th century, the classification of chemical compounds was based on a few physical properties such as color, taste, smell, consistency, solubility, method of preparation, etc. However, in the second half of the 18th century, with the systematic nomenclature of Lavoisier and his collaborators, the chemical composition of compounds emerged as semantically relevant, and subsequently, properties such as color, taste,
smell and consistency were "relegated," so to speak, to being mere indicators of chemical composition, while the latter provided the basis for classification. We have known similar developments with respect to the other terms (see Shapere, 1984;1989 for "electron"; Nersessian, 1984;1987 for "field").

23. Interestingly, model 3 was built on the basis of my preliminary research for the case-study, and later on further refined; for details, see Part IV, Chapter 2.

24. In this respect, Nersessian writes: "[T]he specification and solution of both representational and developmental problems require a method which is 'cognitive' and 'historical'" (1987, pp.163-4, for reasons why).

25. Initially and partially, one may indeed think that a theory of meaning about scientific terms consists in a cluster of empirical hypotheses about the references of those terms. The hypotheses are empirical in the following senses: (i) They may be tested as a linguist or historian or anthropologist might test hypotheses about meaning of terms. The hypothesis must accord with the behavior (in a very, very broad sense) of the scientific user of the term. (ii) They presuppose judgements on the part of the person making those hypotheses as to the truth or falsity of scientific assertions which contain the terms in question. Recall that we are concerned with the meaning of a term contained in given scientific theory. Since we can't blindly assume that all of the assertions containing the term within that theory are true -- after all, scientists make errors, sometimes even gross errors, we must take this fact into account when we formulate empirical meaning hypotheses, and thus "sort out" the errors.

Before formulating more fully and explicitly the criterion of meaning-change underlying my approach, some preliminary set-up seems in order. The questions before us are: (1) Do scientific terms generally change meaning as a result of theory change? [Meaning-Change Question]. (2) And if they do, does that mean that scientific theories are generally incommensurable, and therefore incomparable? [Theory-Comparison Question].

In order to answer the latter question, what we need to determine is whether there was meaning change in a sense of meaning which is relevant to whether statements containing the terms in question deny each other, agree with each other, and so on. We might, on the one hand, give a criterion for judging when a term has the same meaning throughout theory change. This would constitute certainly a first step toward answering (1) the meaning-change
question. But it would not answer (2) the theory-comparison question unless the meaning which it showed to be preserved (or lost) was the relevant kind of meaning. Achinstein [1968] for example offers 'criteria' which answer the meaning-change question negatively. But the kind of meaning (i.e., centrally and semantically relevant properties) that he shows preserved is not the relevant kind, and so his criteria cannot answer (2). In other words, what he offers is not enough. The same point can be made by the way about Fine's account [1967].

On the other hand, we might argue that there is a kind of meaning which is relevant to theory-comparison, and attempt to develop it. But this would not be enough either unless we are given a way of judging that the relevant kind of meaning is preserved in some interesting cases of theory change. As we have seen in Part II, Chapter 1, Scheffler argues that sameness of reference is what is needed for answering (2) the theory-comparison question, and for comparing scientific theories. But he offers no means for determining when reference is the same --or for judging that in some interesting cases reference does turn out to be the same. Field [1973] argues that the traditional notion of reference won't do, and that we need a notion of "partial reference" coupled with a new "definition" (Tarski's style) of truth and falsity. But he judges sameness of partial reference on ad hoc grounds and does not give a criterion (see Chokr, 1990 for details). As for Putnam (Chap.2), we have seen that he offers a causal theory of reference whose main goal is to justify the claim that scientific terms have the same reference despite theory-change. But he overlooks all the interesting cases, where reference may actually change as result of revolutionary theory change, and in the process of developing his theory, makes indefensible assumptions about the "rigidity of reference." Finally, Kitcher proposes the notion of "reference potential" and argues that it will enable us to answer (2) satisfactorily. While I agree that Kitcher's notion is helpful, it serves strictly speaking to answer only (2), and not (1).

In contrast with both these kinds of approaches, the criterion underlying my approach serve to determine when scientific terms have the same 'meaning' and reference through theory change. Thus, in my view, a theory of meaning about scientific terms consists in a cluster of empirical hypotheses about 'meanings' and references of those terms. Unlike Achinstein and other proponents of traditional contextual theories of meaning, I do not believe that all of scientific language is just intensional and that all we need to determine is sameness of meaning, sense, or synonymy. [As Quine has shown (1953), to give a criterion of sameness of meaning presupposes that a
satisfactory analysis of synonymy has first been constructed, which is no simple or easy task, and certainly not one which I intend to undertake herein. Unlike Scheffler and other proponents of reference theories, I don’t believe that all of scientific language is extensional, and that we have no need of any sort of meaning-relations for serious purposes, except what we can get in terms of reference. Much of scientific language undeniably and ineliminably contains expressions which induce intensional contexts. And so, while I agree that reference, as the clearest notion we have, must be included in my criterion (as field of referential possibilities), it also includes considerations of ‘meanings’ i.e., clusters of (correlated and assertible) properties associated on empirical grounds with a given term, which best serve to characterize “meaning as use” (or ‘meaning’ in my terminology, which is distinguished from meaning, as semantic characterization of the things referred to by a term, or Quine’s “meaning”).

Suppose we have two theories T and T’, and a “background context” containing and embodying the best scientific theories of the day --we can only proceed using the best scientific knowledge available, since there is no absolute (first philosophy or methodology) or transcendental (godlike) way to sort out the errors of a theory. In general, this means that the background context will not contain both T and T’ --because T and T’ will often be competing theories, which will not be simultaneously acceptable. This creates a difficulty. Suppose that T is not part of the background context. How is T to be evaluated unless we can first determine that the predicates P’s with which theory T is concerned are the predicates P’ with which the theories of the background context are concerned. If we presuppose that they are, when they are not, this would undermine the grounds for a real evaluation. This problem is especially urgent in cases where we wonder whether T and T’ (or their respective proponent) are concerned with the same things when they use a given term “X”; and where T’ is included in the background context, since it constitutes the best scientific knowledge available on these matters, and must therefore be used in the evaluation and criticism of T (e.g., Stoney’s theory of the electron and Thomson’s or Bohr’s theory of the electron; Dalton molecular theory and Avogadro’s molecular theory). But there are also cases --under similar conditions otherwise, where T and T’ are such that a given term “F” in T has different meanings and is used in many different ways to refer to different things (e.g., F, A), and another term “O” in T’ is used in a way similar to one of the
meanings of "F", to refer to some of the things picked out by "F", but now considered to be neither P nor A, but X (cf: phlogiston/oxygen theory).

The difficulty here in question can be circumvented by a "hermeneutical" procedure, often used by historians, translators, literary critics and philosophers in the phenomenological tradition. We might say that a hypothesis is made that a term has a certain 'meaning' and a certain reference. The hypothesis is tested or defended. Generally, this done by showing that, on the hypothesis, the arguments, associations and assertions of a scientist are as reasonable as one might expect --given previous attributions (on the basis of the principle of humanity) to the scientist in question, of patterns of relations between mental states and the world as similar as possible to ours, and given the scientist's known strengths and weaknesses, and finally given appropriate compensations for his/her mistaken beliefs.

More explicitly, such a procedure might take this form:
(1) A hypothesis is made (implicitly or explicitly) concerning the 'meaning' and the reference of (some key) terms of theory T. Assuming the language of the background context is used as a meta-language, a matrix can be set up giving: (i) associations: key terms $\rightarrow$ clusters of properties. (ii) referential equivalents for terms of T, i.e., the object language.

(2) On the basis of this hypothesis, and given the background context, (i) we seek to correct the associations: terms/clusters of properties; (ii) the assertions of T are assigned truth values, perhaps even 'partial' truth values.

(3) These associations and assignments are defended against the actual associations and assertions given by the author in question (or under the theory in question). At this stage, two fundamental normative principles seems to operate. (A) Principle of Maximization of Truth-Statements: In (most) scientific language use, the aim of the scientist is to make considered assertions from the true class. (B) Principle of Scientific Conscience: A scientist ought to have reasons and even chains of reasons for the associations and assertions made in his/her theories, and for the basic assumptions which underlie these assertions or associations.

(4) Steps (1), (2), (3) are repeated for theory T' (if necessary, i.e., if T' is not part of the background context).

(5) [a] If the cluster of properties associated with a term "X" in T is the same (composition and configuration --relative centrality of properties included) as the one associated with the term "X" in T', then we may say that "X" has the same 'meaning' in both T and T'. Otherwise, we may consider degrees of similarity of 'meaning', or else qualify the possible changes of 'meaning,' and degrees of
'meaning' change, in terms of addition, deletions, and/or restructuration of the cluster of properties associated with the term. If there is some property or cluster of properties C, such that it is judged that the predicates P's of (the corrected) T are the C's, and the P's of the (corrected) T' are the C's, then P's may be considered to have the same reference in the two theories. Otherwise, we will consider that the reference has changed.

The procedure that I propose is more context-sensitive than I have been able to show, and can account for many other cases, i.e., when a term is not used univocally in the theory, or in the background. The schema above can be applied considering one meaning, or different meanings, in different parts of the theory. However, one should note that there will be cases in which there will be no way to decide whether a term in one theory has or has not the same 'meaning', is or is not co-referential with a term (or complex of terms) in another theory. There may also be cases in which sameness of 'meaning' and reference must be determined through a chain of theories more closely connected historically, rather than by considering T and T' directly.

26. I would like to claim what I am proposing here can be also applied to other cases of (revolutionary) scientific development.

27. The list below includes the contributions of those mentioned.

28. A three-stages hypothesis for the replacement of a cluster network by another one seems rather plausible. For example, we may have the following situation:

```
-----(i)  BC1  ---------→  CN1
     |                         /[CN2]
     |                         /  
     BC----(ii) BC2  ---------→  CN1
     |                                           /[CN1]
     |                                           /  
     |----(iii) BC3  ---------→  CN2
```

Given a larger and multivariate background context BC, we can see in (i), how a sub-context BC1 shapes and constrains a given established cluster network CN1; in (ii), how a sub-context BC2, which has evolved from BC1, still shapes and constrains CN1 (frontstage) while at the same time
giving rise to a competing network CN2 (backstage); and finally in (iii), we see how given a different background context, i.e., BC3, CN2 takes "frontstage" and "upstages" CN1, which remains nevertheless on the backstage. Thus, while there is radical change and replacement, it takes place progressively, within a larger background of continuity. And the replaced network does not disappear completely, just as the new network does not (always) appear in Gestalt-like fashion. [Thagard, 1990, p.202 for a schema along similar lines; see also forthcoming discussion of developmental problems].

Developmental Problems

29. I mean (i)-(iv) and not just (i) and (ii), as Nersessian does, and as I will point out in a moment.

30. They are wrongly called incremental or accretion theories, because this presupposes that scientific conceptual change involves only small or minor additions, but not deletions or restructurations.

31. Interestingly, both kinds of theories get their empirical support from different and competing research programs in cognitive psychology.

32. Whenever the notion of GestAlt is invoked, it might be helpful to ask --in order to avoid, among other things, a category-mistake: for whom? The scientist involved in a discovery? Or the scientific community at large? Does the phenomenon of "conversion" to a new network by members of the scientific community take place in gestalt-like fashion, or more progressively, even though ultimately it may entail radical conceptual changes. In passing, it is Kuhn's focus on gestalt-like phenomena (1962/1970) which has led to the emphasis and exclusive attention on radical and discontinuous changes, i.e., to the characterization of "scientific revolutions" in those terms. In his more recent work (1982-1990), Kuhn seems to have changed somewhat his view on this subject, and to have retreated from his exclusive focus (see Part III chap. 3).

33. I am alluding here to various examples which have been the object of concern of philosophers of science in the past two or three decades: e.g., "motion" in Aristotle and Newton, or in Newton and Einstein; "mass" in Newton and Einstein; "field" in Faraday and Einstein; "electron" in Stoney and Bohr, "phlogiston/ oxygen" in Becher/Stahl and Lavoisier, etc.
34. For discussions of problems and issues related to inter-disciplinary transfers, see, to name a few, the works of Tweney, 1981;1984;1985; 1989; Gholson et al, 1989; Gibson, 1985; De Mey, 1982; Nersessian, 1987; 1989; Thagard, 1988;1990; Giere, 1985; 1988; Bloor, 1976; Latour and Woolgar, 1979; Simon, 1977; McCarthy, 1988; Barnes, 1974; McMullin, 1988; Donovan et al, 1988; Longino, 1990, etc.

With respect to one possible avenue for inter-disciplinary transfers, it is worth noting Tweney's comments. He writes: "Our understanding of the principles of human thinking has made enormous strides in recent years. Knowledge about the basic parameters of human memory, learning, language, and problem-solving is now sufficiently thorough to allow discussion of what Anderson calls the 'architecture of cognition'. Our theories now allow us to construct detailed models of 'natural cognition' in very complex domains. However, attempts to use cognitive theory as a basis for the understanding of actual scientific inference are still quite new" (1985, p.190; italics added). Tweney has been interested since the early 80's in what he calls "a cognitive history of science," one which attempts to "reconstruct structural patterns (located in expanding networks of similarly constructed patterns) from historical evidence, and which adopts in order to do so, a conceptual framework based on notions such as schemas, scripts, and heuristics borrowed from cognitive science. In order to ward off any misinterpretation of what he is proposing, Tweney writes: "Let us be absolutely clear about what is proposed. It is not that philosophical considerations should be ignored in formulating normative models of science. Rather, it is that empirical evidence, including the appropriate use of historical evidence, is crucial to certain fundamental problems of philosophical import. What is being proposed is that useful normative models of science will result from the inter-penetration and mutual fertilization of the philosophy of science and the newly emerging sciences of science. The science of the cognitive activity of individual scientists will be one of the keystone sciences of this new joint venture (1981, pp.416-7; italics added). I should say that Tweney's point is well-taken, and I basically agree with it. See also Gibson (1985) for arguments about the "convergence of Kuhn and cognitive psychology" and more specifically, about the application of "schema theory" (Rumelhart, 1980) to the clarification of one Kuhn's notions, i.e., "learned perception of similarity." As for Giere (1988), he argues for a "cognitive approach" in the largest sense possible of 'cognitive' in order to explain science globally. Thagard (1988) proposes a "computational model."
With respect to the other possible avenue for transfers, how not to consider the sociology and history of science, i.e., what we have been learning about the social and historical dimensions of science, without falling into the extremes of the "strong programme" or of historicism. In general, scientists will give the reasons for the assertions of their theories in arguing for their position. They will not generally give reasons for the basic assumptions which are not explicitly asserted, though they are operative, and which are shared by other scientists of their time. Generally however, there will be reasons for them which are available to sociological and historical investigations.

35. See previous note 28.


37. More on these factors in Chapter 2 [case-study].

38. In this way, the new cluster theory would in fact reaffirm the rationality, objectivity and progress of science.

Chapter 2 Case-Study

1. What does this adequacy requirement indicate about the problem of theory-comparison? At least, one thing: there is far more involved in the comparability of scientific theories than the mere continuity of the sort focused upon by Putnam-like reference accounts, i.e., common reference.

2. A point which ought not to be underestimated in our attempt to "regain the past."

3. With the appropriate graphics possibilities on a PC, one could provide a fragment of the Cluster Network involved at various stages of the process.

4. Note in passing that my presentation may suggest quite rightly from a historical point of view that the problem of air and water were 'solved' together, just as the problems of combustion, calcination and respiration were 'solved' together.

5. Incidentally, the fact that metals increase in weight upon calcination was already recorded in Helmont's time (1579-1644), even though Helmont himself did not believe it (Partington, 1961). Later, Boyle (1627-1691) attempted
to explain it in terms of the "fixation of ponderable particles of fire"; but Lavoisier showed this to be experimentally false (cf: experiments on the calcination of tin and lead, using sealed glass retorts, weighed before and after).

6. While 'oxygine' is not yet an element in our sense of the term in 1783, it becomes 'oxygene' (oxygen) in the Nomenclature Chimique of 1787, i.e., an element along with other elements such as light, hydrogen, nitrogen, and caloric.

Chapter 3: Residual Problems

1. Cherniak (1986) for example develops a model of a minimal epistemic agent having fixed limits on cognitive resources, and focuses for this purpose on the "cluster structure" of the deductive inference abilities required of such a minimally rational agent; he claims that this model is more psychologically realistic than classical idealizations.

2. Where should we situate the cluster theory here in question on the continuum between "autonomism" and "historicism" in philosophy of science? What implications can or should we then draw on this basis for philosophy of science, for philosophy in general? To what extent and in what senses does the new cluster theory uphold the tenets of the "new philosophy of science" discussed in the introduction. Possible 'criteria': (i) substantial theoretical claims; (ii) methodology; (iii) underlying or explicit conception of the nature of science, of philosophy of science.

**Autonomism**

\[
\begin{array}{c}
\text{Carnap} \\
\text{Hempel} \\
\text{Scheffler} \\
\text{Putnam} \\
\text{Gasking [?]} \\
\text{Field [?]} \\
\text{Achinstein} \\
\text{Kitcher} \\
\text{Newton-Smith [?]} \\
\end{array}
\]

**Historicism**

\[
\begin{array}{c}
\text{Kuhn [1962]} \\
\text{Feyerabend} \\
\text{Quine [?]} \\
\text{Shapere} \\
\text{Nersessian} \\
\text{Kuhn [1990]} \\
\end{array}
\]
BIBLIOGRAPHY


Athens: Ohio University, pp.84-93.


Dordrecht: Reidel.


Philosophy of Science, vol. 8, Dordrecht: Reidel.


--------, [1976/1929] Considerations on the Doctrine of


Indianapolis: Bobbs-Merrill.


