



National Library  
of Canada

Bibliothèque nationale  
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada  
K1A 0N4

## NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

## AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

**An Examination of the Arguments Against the Naturalistic Paradigm  
of Research in Educational Technology and Their Implications  
for Current Research Practices**

**Steven G. Shaw**

**A Thesis**

**in**

**The Department**

**of**

**Education**

**Presented in Partial Fulfilment of the Requirements  
for the Degree of Doctor of Philosophy at  
Concordia University  
Montreal, Quebec, Canada**

**November 30, 1990**

**© Steven G. Shaw, 1990**



National Library  
of Canada

Bibliothèque nationale  
du Canada

Canadian Theses Service    Service des thèses canadiennes

Ottawa, Canada  
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-64716-7

Canada

## ABSTRACT

An Examination of the Arguments Against the Naturalistic Paradigm  
In Research in Educational Technology  
And Their Implications for Current Research Practices

Steven G. Shaw, Ph.D.  
Concordia University, 1990

The central thesis of this dissertation is that the arguments which have been arraigned against the naturalistic approach in the social-behavioral sciences (and particularly applied fields such as educational technology) are inconclusive. Invariably they invoke false premises and/or invalid inferences. By the "naturalistic" approach, I intend the model of inquiry which is purportedly based on the paradigms encountered in the natural sciences. The basic arguments against naturalism are characterized and evaluated.

While the arguments typically aimed at the "scientific" approach are not persuasive, there are aspects of the current model which must be reformed before we can reasonably expect this approach to succeed. At issue are our methods of theory and hypothesis-testing, and our conceptions of theory and the processes of theory formulation and elaboration. Null-hypothesis testing is inadequate as a methodology for evaluating hypotheses. Our conceptions of theory and theory elaboration are naive, and based on defunct principles of operationism and inductivism. Meehl's well-known criticisms of null-hypothesis testing are summarized and recent attempts to counter these criticisms are weighed and rejected. An alternative approach to theory testing which has been proposed in educational technology, called formative evaluation research, is presented and criticized. Representative examples of misconceptions concerning the nature and function of theory and the process of theory elaboration are drawn from the principal areas of educational technology. These include distance education, instructional design theory and models, the visual literacy movement and realism theory.

The arguments against naturalism are often framed within the context of the debate over the relative merits of "qualitative" approaches to research and the scientific approach. The relationship between these paradigms is addressed, in so far as it relevant to an assessment of

the role of the scientific approach. Relativism, the thesis that qualitative and scientific approaches are equally valid and that the choice between them is based in one's particular values or aesthetics, is rejected. Orthogonalism, the view that they address mutually exclusive sets of questions, is accepted with certain qualifications. Compatibilism, the position which holds that the alternative paradigms are really consistent with one another, is rejected. The principles on which relativism is founded are assessed in detail. These include pragmatism, the Duhem thesis, and conventionalism.

Following the rejection of relativism and compatibilism, the implications for current trends in research practices are drawn. The tendency towards methodological hybridization, in particular, is criticized. If, as argued, qualitative and scientific paradigms are not compatible at a fundamental level (the level which defines such important notions as "truth" and "validity"), then there is no basis on which to propose that they be combined. Furthermore, if the paradigms present different, conflicting, views concerning how knowledge claims are justified, then it makes little sense to attempt to support propositions advanced under the one approach with the results obtained under the second, as occurs in the strategy called methodological triangulation. Given their different logics of justification, claims advanced under one paradigm do not carry evidentiary weight when viewed from the perspective of the other.

### Acknowledgements

I would like to acknowledge several debts in connection with this dissertation and the course of studies leading up to it. First, and foremost, I would like to thank my supervisor, Dr. Robert Bernard, for the interest he has shown in my work, and for the encouragement and assistance he has provided both with regard to this dissertation and concerning the various projects I have undertaken outside the confines of the program and Concordia University.

I would also like to thank my internal committee members, Dr. Richard Schmid and Dr. Dennis Dicks, for their careful scrutiny of various drafts including some rather rough early manuscripts, and for several helpful suggestions regarding references, points of presentation and clarification of substantive issues. Like Dr. Bernard, they have been generous with their time and thoughts and have also allowed me the opportunity to participate in several of their own projects. Dr. Dicks graciously agreed to continue as a committee member during his sabbatical leave.

Dr. P. D. Mitchell and Dr. G. Boyd of the Department of Education were involved in the supervision of my doctoral studies at an earlier stage. Their interest in philosophical issues, both methodological and ethical, has been instrumental in providing a unique flavour to Concordia's Educational Technology programs and in creating the atmosphere in which it was possible to pursue the interests reflected in this work.

A note of thanks, also, to my external examiner, Dr. W. Winn, whose ruminations on the theoretical directions taken in the field I have enjoyed reading in published form. It was in fact a public lecture delivered at Concordia by Dr. Winn in 1988 that provided some of my inspiration for writing this thesis.

I reserve the blame for any omissions or errors which may remain.

## TABLE OF CONTENTS

<b>CHAPTER 1: Preface</b>	1
Thesis	1
Context	1
Overview	3
Some Preliminary Definitions	6
Technology	6
Educational Technology	9
Science	15
Paradigms	16
<b>CHAPTER 2: External Criticisms of the Naturalistic Approach</b>	18
The Arguments	20
1. The "Value-Ladenness" of Social Science Discourse	22
2. The Argument From the Principle of Uncertainty	25
3. The Impossibility of Formulating Reliable Generalizations in Social Scientific Inquiry	27
4. The Sources of Scientific Knowledge	31
5. The Irrelevance of Theoretical Inquiry to Practical Educational Problems	33
6. The Intractable Complexity of Social Phenomena	35
7. The Multiplicity of Social Realities	37
8. The Superiority of "Grounded" Theory	41
Prospects for Generalizable Non-Naturalistic Theorizing	
9. Hermeneutic versus Deductive Modes of Reasoning	45
Conclusion	50
<b>CHAPTER 3: Internal Criticisms of the Naturalistic Approach</b>	53
The Defects of Operationism	53
The Requirement of Theoreticity	55

Routing Out False Conjectures: The Limitations of Social Science Methodology	62
The Formative Evaluation Approach to Theory Testing	70
Conclusion	73
CHAPTER 4: Misconceptions Concerning Theory and Theory Development: Some Illustrations from the Literature	76
Holmberg's Theoretical Framework for Mediated Instruction	77
Deductive versus Inductive Theorizing	77
Descriptive versus Prescriptive Theory	79
Metaphor as Theory	80
Common Sense Knowledge as Theory	81
Tautologous Propositions as Theory	82
The Visual Literacy Movement: Basing a Research Program on Metaphor	83
Theories of Instruction and Instructional Design	87
The Relations Among Theories and Models	89
Instructional Theory, Learning Theory and ID Models	91
Richey's Prescriptions for ID Theory Development	92
Summary	96
Further Examples	98
Readjusting the Goals of Research: Low Level Generalization versus Deeper Explanation	100
Educational Psychology as Conceptual Analysis	101
Conclusion	103
CHAPTER 5: The Relationship Among Alternative Paradigms	105
The Alternatives	106
Incompatibilism	106
Orthogonalism	107
Compatibilism	107
Relativism	108
Compatibilism	109



Pragmatism as a Basis for Compatibilism	109
Denial of the Fact-Value Distinction as a Basis for Compatibilism	113
Summary of the Case for Compatibilism	114
The Ambiguity of the Term "Methodology"	115
Parallel Accounts of Criteria for Judging Qualitative and Quantitative Inquiry	116
Compatibilism at Various Levels	117
The Case for Combining Methodologies	119
Orthogonalism	120
The Arguments for Relativism	124
The Problem of Confirmation	125
The Underdetermination of Theory by Logic	126
The Underdetermination of Theory by Experience	127
The Duhem Thesis, Incommensurability and Global Conventionalism	128
Initial Conceptualization of GC	128
The Duhem Thesis	130
The Weltanschauungen View: The Incommensurability of Paradigms	135
GC and Criteria of Theory Synonymy	140
Some Conceptual Confusions	149
Empirical Equivalence	152
Craig's Theorem	154
The Lowenheim-Skolem Theorem	155
A Final Argument for Relativism: Values as the Basis for Theory Selection	158
Summary of the Case for Relativism	160
Conclusion	161
CHAPTER 6: Concluding Remarks	165
FOONOTES	174
REFERENCES	177

## CHAPTER 1

### Preface

#### *Thesis*

This thesis concerns the "scientific" approach to social-behavioural inquiry and its role in the field of research and practice in educational technology (henceforth, ET). It will be argued that current criticisms of scientific inquiry in the domains of education and instruction are insufficient to establish that such an approach is illegitimate, that it cannot succeed.

At the same time, it will also be contended that the current model or paradigm of scientific research suffers serious defects -- relating to the logic of hypothesis testing, and the model for the structure and semantics of theories which has been adopted. These flaws are, if you like, conditionally fatal. That is to say that so long as they continue to characterize our scientific model of research it cannot accomplish its purposes. However, it appears the difficulties involved can be addressed. I will argue that remedies are available which, though relatively drastic, leave intact the basic core concepts and assumptions of the scientific approach.

#### *Context*

The context and rationale for the topic is provided by the debate which has emerged over the last 15 years concerning what constitutes an appropriate methodology for acquiring an understanding of educational or instructional phenomena and, by extension, dealing with problems related to educational systems. Of course, concern over the generalizability of the scientific approach, a mode of inquiry which evolved in the domain of physical systems, to human behaviour has a much longer history; but the debate in the educational field has become more intense in recent years, presumably owing both to concern over the growing magnitude of problems in public and higher education, and (of particular relevance to the subfield of ET) an increasing requirement for cost-effective training in the private sector, public service and the military. These concerns have led to recent widespread acceptance of a variety of new methodological approaches, and to more vigorous criticism of the traditional scientific approach, as it is utilized both in basic research and in evaluation studies, as ineffective. Current debate in the arena of educational research usually arraigns the assumptions, procedures and techniques

of the traditional scientific framework against those shared by a variety of approaches which derive from the practices of sociology and anthropology, and which are subsumed under the rubric of "qualitative" inquiry.

In the more narrow subfield of ET there is, in addition, another dimension to the present methodological controversy. There has been considerable argumentation concerning the relative merits of traditional science versus a variety of models of inquiry which are, in several instances, quite distinct in their basic premises, but which are also commonly lumped under one heading, namely, "systems approaches". ET is perhaps unique in this respect as other areas such as public policy studies and management theory have recently shifted away from any significant reliance on systems modelling and thinking.

This dissertation will focus on the controversy as it is framed in terms of the debate concerning the merits of the qualitative versus the scientific modes of inquiry. The reasons for this particular focus are threefold. First, the relationship between qualitative and scientific inquiry is more interesting and less well understood than the relationship between traditional science and systems approaches. Secondly, as remarked the methodological debate in the general literature of education has concerned primarily the qualitative and scientific approaches. The field of ET has, not surprisingly, begun to follow suit. Finally, while support for systems approaches is waning in areas of applied social science, support for qualitative methods continues to acquire momentum.

There is an important terminological issue which must be addressed straight away. In the philosophical literature the term "naturalistic" is understood to mean the application of the approach familiar to us from the natural sciences to the social/behavioural domain. Conversely, in the literature of education and educational technology the term has been appropriated to cover virtually any approach that is qualitative and interpretive. This may include, for example, the case study, action research, ethnography, and ethnomethodology and the various subcategories of these. The convention followed in this dissertation is that the term "naturalism" will be used to indicate the approach that is said to be based on the natural sciences, while the term "qualitative" will generally serve as a convenient label for the approaches of

ethnmethodology, ethnography, phenomenologically oriented inquiry and so forth.

### *Overview*

Criticisms of the scientific approach may be labelled as external or internal. External criticisms comprise objections to any of the fundamental assumptions or procedures of the scientific method raised by proponents of alternative frameworks. Internal criticisms, posed by practitioners of the scientific paradigm, typically are less sweeping in character, involving more specific emendations regarding procedures, techniques or instrumentation. Chapter 2 will present an overview of external criticisms current in the literature. These criticisms can also be categorized as potentially "fatal". By this I intend that they speak to the most basic assumptions of the scientific paradigm regarding e.g., the nature of explanation or the fundamental characteristics of social-behavioural phenomena. If the arguments for any of these criticisms go through, then the scientific paradigm is thereby shown to be unworkable and unsalvageable.

If these arguments are not cogent, as I shall argue in Chapter 2, then (putting aside for the moment the defects alluded to above, which I maintain are addressable) it may be concluded that, at present, there are no decisive reasons for rejecting the scientific approach. Admittedly, a case might still be made, perhaps, that there are no compelling positive reasons for advocating this particular mode of inquiry. Such a line of reasoning would take as its point of departure the failure of the scientific approach to achieve significant results in terms of the cumulative development of theory. The promise of such cumulative development is, after all, the chief rationale or justification for utilizing the methodology of science.

Even granting this relative failure as a premise, it will be seen that it is still possible to defuse this line of reasoning. There are credible reasons for not attaching any significant weight to the lack of accomplishments attributable to the scientific approach to date. These include limitations of the standard analytical tools employed, approaches taken to theory elaboration and testing, and certain sociological aspects of soft sciences research programmes -- the existence of which do not necessarily impugn the basic project of science.

Such problems constitute the class of internal criticisms mentioned above. These are addressed in Chapter 3, with emphases on the following:

(1). In the social/behavioural domain, the received view of the semantics of scientific theories is influenced by operationalism, which is a defunct philosophy of science. Against operationalism, which holds that all terms in a theory must be explicitly defined in terms of procedures of measurement and experimental manipulations, I shall argue that scientific theories are characterized by the property of theoreticity. By "theoreticity" I mean the possession of certain terms designating hypothetical or unobservable constructs, some of which are only implicitly defined.

(2). The logic of theory testing, which focuses on the refutation of the null hypothesis rather than a precise experimental prediction, is too weak to furnish the basis for a methodology which could be described as "self-correcting". It also discourages the development of more sophisticated, more precise, theories. Meehl's well-known criticisms of null hypothesis testing will be presented, and a recent attempt to answer these objections will be assessed and rebutted.

(3). Lastly, a strategy for reforming "scientific" research which has recently been advocated in the field of ET will be described and evaluated. I am referring to what Reigeluth (1989) now calls "formative evaluation" research and what Clark (1989), following Reigeluth's earlier (1983) terminology, has labelled "development" research. I will reject this form of research as inherently incapable of providing any genuine, unambiguous test of a theory, or any component of a theory.

Chapter 4 will continue on a more concrete plane the general theme of Chapter 3 concerning the inadequacies of established conceptions of theory and theory testing in the field. In this portion of the dissertation I will examine some representative cases of confusions concerning the nature of theory and the process of theory-testing, drawn from the core areas and literature of educational technology. Examples will be presented of a variety of discursive objects -- metaphors, taxonomies, hierarchies, tautologies and low-level empirical generalizations -- which have been passed off as theory. The examples selected are taken from the following areas: instructional theory and instructional design models, theories of visual learning or the visual literacy movement, realism theory, and theories of distance education or mediated

instruction. Conceptions of the process by which theories are developed and refined will also be examined.

Chapter 5 will address the complex issue of the relationship between the qualitative and scientific paradigms. The questions which are raised are the following: Are these approaches "orthogonal" (do they have different purposes and do they only deal, appropriately, with different sets of issues)? Are they compatible (and in what different senses, or at what different levels, are they compatible or incompatible)? And, is relativism, the position that these approaches are equally functional and equally good and that the choice between them is cognitively-speaking an arbitrary one, valid?

The relationship of these questions to our general theme developed through Chapters 3 and 4 is this. Relativism, compatibilism, and non-orthogonalism have acquired a considerable following. These philosophical positions in turn provide the basis or legitimation for emerging methodological trends aimed at combining elements of both qualitative and quantitative methodology to address the same issues more effectively. These trends include the use of composite methodology, which I call "hybridization", and methodological triangulation. Such trends serve to further dilute methodology (and the very concept of a *method* of inquiry) and to render the task of refining and reforming our methods of inquiry even more problematic. They are counter-productive; they simply exacerbate the problems exposed in Chapters 3 and 4.

Addressing the issue of relativism will require an evaluation of post-positivistic epistemological principles which furnish the usual justifications for this position. These include incommensurability (the view that rival theories or paradigms cannot be compared) and conventionalism (the view that theories are arbitrary constructions and that there may exist alternative theories which are equally confirmed by any possible evidence). Conventionalism has been associated with a variety of other philosophical viewpoints, including: the Duhem thesis, Quine's brand of pragmatism, instrumentalism, and a brand of relativism inspired by the Kuhn-Feyerabend view of science and its progress.

The relationships between these additional philosophical theses and conventionalism will be clarified. On closer scrutiny it will be seen that they are quite distinct. An innovative approach

to conceptualizing the whole issue of conventionalism will also be presented. The key will be the introduction, and defence, of a new formal characterization of theory synonymy elaborated using the tools of model theory. Armed with this characterization it will be shown that the case for the underdetermination of theory by evidence is much weaker than has been generally supposed. While it will not be demonstrated that the thesis is false, it will be established that there is no sound case for it. Moreover, under the rigorous analysis that will be presented I believe conventionalism appears counter-intuitive.

The case against incommensurability, the other leg on which relativism may stand, is stronger. Incommensurability appears to be demonstrably false or, worse, incoherent.

Thus, Chapter 5 will culminate in a rejection of relativism. Along with a repudiation of compatibilism and a qualified acceptance of orthogonalism, this will eliminate the standard philosophical justifications for compounding disparate methodologies and complete our critique of current conceptions of theory, theory-testing and theory development in the field of ET.

Chapter 6 will serve to present certain minor points left unattended in the earlier chapters and to synthesize the conclusions reached once more. The impression left by Chapters 2 through 5, given their critical orientation, may be one of a wholly negative view of the field. This will be corrected in the concluding chapter with a brief discussion of some of the more positive developments and directions in theory development and testing in ET. These advances are not, however, to be construed as mitigating the basic arguments and worries raised in the preceding chapters.

#### *Some Preliminary Definitions*

Before proceeding, we require some working definitions for terms such as "technology", "educational technology", "paradigm" and "science":

##### *Technology*

There are several alternative conceptions of technology. One popular definition makes technology synonymous with "applied science". This conceptualization generally goes hand in hand with a particular hierarchical view of the relationship between basic or "pure" science and technology according to which advances in basic research lead to technological developments,

but not vice versa. It also frequently associates technology very closely, even criterially, with "hardware" (i.e., with machines or physical systems of one kind or another). Berry provides a typical expression of the definition of technology as applied science: "any application of the discoveries of science, or the SCIENTIFIC METHOD, to the problems of man and his environment . . ." (Berry, 1977).

While this analysis has enjoyed considerable popularity in the literature of the philosophy and history of technology in the twentieth century -- it might even be called the received view -- it has recently lost ground to other conceptualizations and their attendant positions regarding the diffusion of technology. Most notably, it now seems that historians of science and technology are more apt to recognize that frequently advances in science are motivated by developments in technology, rather than the reverse (de Brisson, 1987). The development of the combustion engine, for example, lead to some major theoretical advances in the field of thermodynamics by the French physicist Carnot (1824). It is probably safe to say that anyone currently writing about the diffusion of technology recognizes the bilateral nature of the relationship between pure science and technology.

Even more important, perhaps, is an alternative analysis of the notion of technology which simply does not equate technology either with applied science or with hardware. This analysis can be supported on two grounds. To begin with, the phenomena of applied science is historically a rather recent one. Applied science emerged only with the advent of the industrial revolution and it is uniquely a product of Western Civilization. So, from a broad historical perspective, any conceptualization of technology which reduces it to applied science must appear somewhat parochial. Second, and more relevant to our purposes, is the consideration that a conceptualization of technology which does not restrict the scope of the term to either systems comprising hardware or to the application of scientific knowledge or methods to the solution of practical problems is richer and more interesting from the point of view of any discussion of ET and its goals and tools.

For one, it does not beg the question of the relevance of scientific inquiry to the field, it allows us to broach the question whether effective practice may be based on modes of thought,



or forms of knowledge, other than those encompassed by traditional science -- qualitative inquiry, systems thinking or systems theory, for example. For another, it does not force us to accommodate a narrow equipment-oriented conceptualization of ET such as that which dominated in the field in the early days of the audio-visual aids movement (Davies, 1978). Such a restricted understanding of ET would certainly exclude major areas of endeavour within the field, such as Instructional Design. Moreover, it would also have the effect of reducing the educational technologist to the status of a technician (one who uses technologies), rather than a technologist (one who develops technologies), since by and large those of us in the field utilize, but do not invent, various types of mechanical or electronic hardware.

For our purposes, then, it seems best to adopt the widest or weakest conception of technology, and we may do well to embrace the definition of "technique" which Ellul (1964) provides in his classic work concerning technological society: "technique is the totality of methods rationally arrived at and having absolute efficiency (for a given stage of development) in every field of human activity" (p. xxv).

Ellul identifies five subdivisions of technique -- mechanical, intellectual, economic, organizational, human (any domain in which man becomes himself the object of technique such as, e.g., medicine, genetics, propaganda, pedagogical techniques, publicity) -- which can be mapped quite neatly on to the various concerns and subdivisions within the field of ET. In various combinations, his five categories of technique can easily subsume, for example, the sub-areas of ET identified by Mitchell (1972), namely: educational psychotechnology, educational management technology, educational systems technology (the planning, design, construction and evaluation of educational systems), and educational planning technology (educational systems technology at a more global level).

This, of course, assumes that technologies do exist in these ET areas and that they are not predominately craft-based. If, refining Ellul's contribution somewhat, we define a technology or technique as a problem-solving approach which has the characteristics of being (a) generalizable to a class of related situations, (b) effective (leads usually to a solution) and (c) efficient (leads to a solution with an acceptable expenditure of resources), then it is also an open

question whether we truly do already possess such a thing in certain enterprises in ET, or whether our endeavours are essentially craft-based (cf. Clark, 1984, for an expression of concern about this possibility). It is clear that we do not possess in the area of ISD (instructional systems design), for example, a technology in the manner in which engineering principles and knowledge constitute, in Ellul's sense, a technology for the design of material systems.

### *Educational Technology*

The concept of ET has changed over time as the scope and aims of the field have evolved. Historically, the roots of the field are to be found in the audio-visual aids movement and the activities of organizations such as the Department of Audio-Visual-Instruction (DAVI-NEA) which was active during the late fifties. As mentioned above, the phrase "educational technology" had a relatively narrow, machine-oriented connotation at this time. DAVI subsequently became the Association for Educational Communications and Technology (AECT) which grew into perhaps the major ET affiliation through the sixties and seventies. The title of this organization reflected the growth of ET into a field concerned with all aspects of the design, implementation, management and evaluation of educational or instructional systems. Understood in this wider sense, ET potentially draws on any of a wide variety of disciplines and areas of study: organizational theory; administration; human resources planning; communications; educational psychology, including learning theory (drawing from behavioural, cognitive, information-processing oriented, and developmental psychology) and educational measurement and evaluation; artificial intelligence or cognitive science studies; instructional technology (instructional "theory" and models of instructional design), the study of human factors or man-machine interaction and; educational cybernetics -- just to name the ones which come immediately to mind.

Mitchell (1972) furnishes a definition which captures the inclusiveness of the field as it has evolved during the last two and a half decades: "Educational technology is an area of study and practice (within education) concerned with all aspects of the organisation of educational systems and procedures whereby resources are allocated to achieve specified and potentially replicable educational outcomes" (p. 325).

Mitchell's definition is echoed by that which the Association for Educational Communications and Technology (henceforth AECT) itself published at about the same time, which delineates the field as one involved with "the facilitation of human learning through the systematic identification, development, organization, and utilization of a full range of learning resources, and through the management of these processes" (1972, p. 36).

Both Mitchell and the AECT are at pains to distinguish ET from other areas of study within education, such as educational psychology, and to show that ET is not coextensive with the entire field of education – an impression which the foregoing definitions might create. AECT argues that ET is a movement within the field of education based on a certain approach and a certain philosophy. The approach identified is technological, and here technology is defined, in a narrower sense than I have chosen, as "the systematic application of scientific and other organized knowledge to practical tasks" (p. 36). This in itself hardly serves to distinguish ET from other areas of the field of education, except perhaps to the extent that it conveys an orientation towards practical matters or problem-solving. This might serve to distinguish it from say, educational psychology, though it might only involve a matter of emphasis or degree.

Perhaps the three patterns of interest which are identified as forming the philosophical orientation of ET are intended to be more definitive of the field. AECT maintains that these three strands, which historically emerged separately, have been synthesized into an identifiable "total approach" and that this approach secures the uniqueness of the field while also providing its rationale. They are: (1) a dependence on a wide range of resources for learning; (2) an emphasis on individualized and personalized learning, and; (3) a reliance on the systems approach (p. 37).

Again, are such considerations sufficient to distinguish ET? (1) and (2) might well serve as the foci for a program of research or studies in educational psychology, and it is simply a contingent fact that ET has tended, historically, to move in these directions more so than educational psychology or, say, the field of curriculum and instruction.

This leaves the systems approach as a potential defining characteristic. If we mean by the systems approach, the decision to regard systems as whole entities, to take the perspective

of the "big picture", if you like, then clearly there is nothing to preclude thinkers or practitioners in other subfields of education from doing the same. If, on the other hand, the "systems approach" is taken to designate the use of a particular systems approach such as cybernetics or General Systems Theory (henceforth, GST), then this would certainly serve to isolate ET from other areas. This is apparently the position suggested by Brahm (1973) who, after offering another sweeping description of the field which includes educational cybernetics, concludes his discussion by remarking that we are possessed of a special tool for acquiring information and knowledge and for its integration and application in an effective technology of education: "Fortunately, we now have an approach that is invaluable, that of General Systems Theory, which concerns locating isomorphic principles that have descriptive and normative applications" (p. 76).

Unfortunately, this, too, fails to establish any rigid boundary to demarcate ET. The simple reason for this is that we do not employ the methods of GST. There is no single instance of a significant application of the technique of identifying isomorphic principles in the field of ET. And, leaving aside the question of the value or validity of this technique, the reason is not hard to find. As Kerr (1989) remarks, we simply have not acquired anything resembling the knowledge base concerning the domain of education or ET that would allow us to benefit from a general systems approach to the "important educational problems of our time" (p. 145).

It remains to be seen whether Mitchell is able to establish that ET, as mapped out by him, can be shown to be unique:

Though it may seem that educational technology has been made coextensive with education, this is not the case. First, the concept of education is not clear; secondly, educational technology so defined is not directly concerned with all aspects of education (e.g., philosophy or sociology of education, administration, ordinary teaching). (1972, p. 325)

As to the first part, it is hardly obvious that the fuzziness of the concept of education is any help in delimiting a unique subfield of ET. As for the second part, I suspect that this distinction may have provided a boundary twenty years ago; however, it has since dissolved or is at least rapidly

dissolving. Educational technologists have certainly tried to have an impact on traditional teaching, both at the public school level and in higher education, through their efforts to influence teacher training and faculty development strategies and practices – though the extent of their influence admittedly seems limited so far (cf. Reiser, 1988; Bratton, 1988; Schiffman & Gansneder, 1988; Gustafson & Bratton, 1984). They have also concerned themselves explicitly with administration. For example, they have played some role in the adoption of strategic planning in university administrations. With the increasing utilization of qualitative approaches the sociological aspects of educational systems and their import for ET approaches have also become legitimate topics.

Finally, it is clear that ET must, as a field, be highly sensitized to philosophical issues. It is implausible that ET approaches or "fixes" will be accepted, unless they can be legitimized within the framework of a palatable philosophy of education. A number of commentators on the field of ET, including Mitchell himself in many of his writings, have emphasized this last point. Hooper, for example, states that "the strength of educational technology will ultimately depend on the quality of the philosophy and the validity of the science of learning that undergirds it" (1971, p. 139).

It is obvious, I think, that some of the resistance to ET solutions is due to the perception that the goal of ET is to replace teachers with machine mediated instruction, and at least part of that resistance is ideological in nature. But, again, it does not appear that ET is unified by any overarching normative precepts concerning this issue. Some major figures clearly do espouse the view that human teachers can be replaced, or their role diminished, through ET and that such a displacement would generally be a good thing. This is a position taken quite explicitly by Heinich (1984). The article in question was challenged in a reply by Clark (1984), but it is interesting to note that Clark's objection is really that ET does not possess the technology to accomplish this; he does not quibble with the program Heinich advocates on any philosophical or normative grounds.

Of course, many educational technologists do maintain that their goal is to supplement or enhance what a human tutor can accomplish, and to provide the best possible alternative

where none is available. Indeed, for many practitioners the justification and rationale for ET is that instruction can be further humanized and democratized through the judicious application of systems and components designed, developed and implemented according to principles of the field.

It seems, then, that it is difficult to establish a unique identity for ET on the grounds either of any philosophical or any methodological principles universally subscribed to by researchers and practitioners of the field. Moreover, the recent proliferation of methodological orientations has made it even more difficult to encompass all the activities and viewpoints under the umbrella of a single definition. The variety within the field is reflected in the different emphases of the curricula in graduate programs in ET, and also in the numerous professional associations which have emerged with ties to the field that exhibit different memberships and serve different interests: the National Society for Performance and Instruction, the Association for Development of Computer-Based Instruction, the Society for Applied Learning Technology, the Human Factors Society, the Association for Educational Communications and Technology, the American Society for Training and Development, the Association for Media and Technology in Education in Canada, the Society for Professors of Educational Technology, to name just some existing organizations. There is also evidence that elements of ET are increasingly being incorporated into the repertoire of other subfields. Reiser (1988) concludes that basic text in educational psychology, such as Good and Brophy (1986) and Woolfolk (1987), are increasingly giving attention to various instructional design principles. And, even more striking, House and Bratton (1987) detect an emerging pattern of convergence in their study of longitudinal curriculum: changes in instructional technology and educational psychology doctoral programs

Neither of these circumstances – the failure of the field to coalesce around some basic shared set of philosophical and methodological precepts, and the appropriation of elements of the *modus operandi* of ET by other fields – need necessarily be viewed with too much concern. Hannafin (1989), for instance, takes a positive view of our diversity. While recognizing that our plurality impedes the development of ET as a discipline or a profession, he argues that "the unification of focus would weaken the breadth of our foundation and limit our capacity to

advance the 'new best way'," and that "the liabilities of adherence to a unitary view of the field more than outweigh the potential advantages" (p. 141). Likewise, the appearance of elements of our accumulated knowledge in other territories should be regarded as a legitimation of ET.

In terms of providing the framework for this present work, there is a final division within ET which must be noted. This is the distinction between what I shall call "radical" ET (after Duchastel, 1989) and "routine" ET. Radical ET concerns itself with the possibility of effecting major changes in society through the transformation of educational systems and practices. Radical ET involves the notion of change agency, of introducing significant alterations in the values, organization, and processes of institutionalized or formal education. This is a "political position" in the broad sense of the term. It necessitates winning over, infiltrating, or otherwise influencing, a lattice of social institutions involving government and institutions of public, private, and higher education -- institutions which often possess an ethos that is overtly hostile to ET approaches and solutions. In contrast to the political character of radical ET (Duchastel labels it "idealistic" and "elitist" rather than "political"), routine ET concerns the application of ET approaches (particular instructional systems design or ISD) in environments -- principally public service, the military, and corporate training -- where their benefits have become clear and where they are consequently well-received and well-entrenched.

I take it that the role of scientific research is to provide us with a better understanding of instructional phenomena and hence, possibly, also a better control over the events of instruction and their outcomes. This role has a crucial place in both radical and routine ET; hence for the purposes of this dissertation it is not critical that I align myself with either position. In truth, I believe my orientation lies somewhere between the two. I do not believe that ET should ever be expected to significantly transform society through its own activities alone, but I do believe that ET has a role to play as an agent of change in transforming our present educational system into something more effective and efficient than the status quo. However, the arguments for that viewpoint lie outside the scope of this present work.

In general, then, I would hold that the differences between ET and other subfields of education, such as curriculum and instruction or educational psychology, is presently a matter of

the specific emphases placed on certain themes (e.g., learning rather than teaching), and a focus on certain contexts rather than others (training and mediated learning, rather than traditional forms of public and higher education, for example).

I do not believe that we require a "model" for ET in the sense in which some have argued for revising and consolidating our problem-solving activities along the lines of an "engineering" model, an "applied science" model, or the model of "medical research". The search for such models is symptomatic of the widespread dissatisfaction with the results of research efforts in the field. However, the arguments for a monolithic approach that will serve to mark ET as something distinct within the boundaries of education, are counterproductive. ET is a field of study itself, and it encompasses a wide range of issues and contexts. It is likely that such a broad scope of activities will require flexibility and a range of tools. My defence of the methodology of science is not part of an effort to establish this methodology as the foundation of the field, though I would argue that it is uniquely suited to a certain range of purposes, relating to the search for "replicable outcomes", that are quite central to the field.

### *Science*

In general we can say that science is a mode of inquiry whose discourse comprises theories. Theories, in the scientific sense, have the following general characteristics. They describe some aspect or portion of reality; they provide an explanation of the behaviour exhibited by that portion of reality which they describe; they have predictive power (that is to say, they enable us to make predictions concerning the reality they describe and explain), they are a richly interconnected, highly systemic form of discourse; they are parsimonious or elegant, they are empirically falsifiable, and; they have scope (that is, they are generally able to subsume and explain a number of seemingly diverse phenomena). Scientific method, based on experimentation, is a self-correcting methodology that advances theoretical discourse through the process of posing theoretical conjectures and then eliminating false hypotheses.

This is perhaps all that we can say about theories in general. However, we can make detailed specifications of the logical structure and semantics of particular theories or even of theories belonging to a particular class (say, relativistic theories of space and time) or to a



particular domain (e.g., theories of physics in general) (Beth, 1961). We can also enhance our understanding of science by elaborating such general issues as the logic underlying the validation of theories (the "logic of justification"), the nature of measurement, the relationship between explanation and prediction, and the nature and role of theoretical terms, definitions and conventions in theories. At the beginning of Chapter 2, a more detailed description of scientific method in the social-behavioural domain will be provided, one which addresses the semantics and structure of social-behavioural theories, the analytical tools used in their formulation, and the logic of justification that is employed.

### *Paradigms*

The term "paradigm" was originated by Kuhn (1962) in his work entitled *The Structure of Scientific Revolutions*, one of the most influential post-positivistic works in the philosophy of science, and certainly one of the most widely read by social scientists. Kuhn's commentators rightly criticized him for the ambiguity he instilled in the notion of a paradigm. Masterman (1970), for example, identified 21 different senses of the term, falling into three groups (metaphysical, sociological, and construct or artifact paradigms). However, the central meaning seems to fall within the last of these three categories, and it pertains to a model from which a tradition may be abstracted or derived: "some accepted examples of actual scientific practice -- examples which include law, theory, application and instrumentation together -- provide models from which spring particular coherent traditions of scientific research" (Kuhn, 1962, pp. 10-11).

The term is used ambiguously in the literature dealing with methodology or foundational issues in the social sciences to designate either the model from which the prescriptive guidelines for carrying out scientific research are abstracted, or for those abstracted principles themselves. In the latter sense the notion of a paradigm fits into Lakatos' (1970) concept of a research programme, and is virtually synonymous with Laudan's (1978) research traditions:

A research tradition is a set of general assumptions about the entities and processes in a domain of study and about the appropriate methods to be used for investigating the problems and constructing the theories in that domain. (Laudan, 1978, p. 81)

In general, the ambiguity of the term "paradigm" is not harmful, but there are at least two

circumstances where it may do considerable damage. The first is where we are dealing with a domain where research traditions have not been abstracted from model cases, simply because no such model cases exist. This is the case in the social and behavioural sciences, where principles of research are based largely on an abstract philosophical consideration of the foundations of knowledge and the nature of understanding (Mackenzie, 1977). The use of the term "paradigm" here is misleading to the extent that it may obscure the fact that research practices are not based on model cases (artifacts) of a highly successful nature.

The second is where we investigate questions of incommensurability and relativity. Kuhn has argued that rival paradigms are incommensurable. His arguments, as we shall see later, are based largely on the notion that terms acquire their meaning within the context of a theory. It follows from this premise that sentences in one theory cannot be compared with those of another simply because there is no assurance that the same terms, appearing in two different theories, carry the same meanings.

Given Kuhn's premise, it makes sense to talk about the incommensurability of paradigms in the sense of models, for these models are theories. Kuhn's claim, though, is that the empirical claims of two theories cannot be contrasted. If, on the other hand, we construe "paradigm" as referring to the methodological principles abstracted from a model theory, then it is not clear that rival paradigms are incommensurable. First of all, such abstract principles are not expressed within the object language of any theory. And, secondly, they are not empirical in nature.

Thus, when we come to discuss the issues of relativism and incommensurability of rival paradigms, we will have to be careful which sense of the term we are employing in our reasoning. However, in general, I will employ the term in the sense closest to Laudan's concept of a research tradition.

## CHAPTER 2

### External Criticisms of the Naturalistic Approach

The purpose of this chapter is to assess the arguments against the naturalistic approach to social science. Before proceeding to this task, however, there is a preliminary item which must be addressed. We need to define more precisely what is meant by the naturalistic approach in social/behavioural science. Just as the term "qualitative" as we are now employing it masks the diversity among the approaches which it subsumes, so it is also the case that several different modes of inquiry purport to be naturalistic. By naturalistic I shall intend the model of inquiry which features the following basic characteristics:

(1) It must provide *causal* explanations of the phenomena to which it is applied. This allows for experimental and quasi-experimental approaches, but excludes *ex post facto* studies and correlational research. This is not to say that such approaches do not yield insights which are important to those seeking causal theories, but rather only that generalizations that are purely correlational do not constitute theories or laws in the sense required by the natural sciences. It is not easy to define exactly what constitutes a theory in the sense of the natural sciences but, minimally, there are these requirements: they are verifiable, falsifiable, descriptive of the phenomena, explanatory (in a causal sense), and have predictive power. Correlations have predictive power and are certainly descriptive. But they are not explanatory (a statement expressing a certain correlation does not account for that relation), and in a certain sense they are not falsified but rather are modified by any subsequent data.

For the most part, certainly in the subfield of educational technology, research is quasi-experimental rather than experimental. This is not necessarily a point of contention. A strong argument can be made to favour natural rather than laboratory settings for research; after all, these are the environments in which instructional artifacts and approaches must succeed. And the use of natural settings for research, with their usual logistical and bureaucratic constraints, often mitigates in favour of quasi-experimental design involving the random assignment of intact groups to different treatment conditions, rather than true random selection wherein each subject has an equal likelihood of being in a given treatment. However, as we shall see, this prejudice exacerbates problems associated with the logic employed in testing hypotheses in this model.

(2) There is the assumption that the phenomena can be measured or quantified, where variables are not simply dichotomous. A further, related, assumption is that the instrumentation used to measure can be standardized.

(3) Further characteristics of the naturalistic approach as it has developed in the applied fields of education and educational technology, and in the social/behavioural sciences in general, include a reliance on statistical models (especially multivariate analysis of variance and regression analysis, and their derivatives) and the logic of null hypothesis testing.

(4) Finally, there is a strong tendency towards operationalism as an underlying philosophy of science.

Note that (1) and (2) are crucial or defining components of a naturalistic approach. (3) and (4), on the other hand, are contingent. They are characteristics which have emerged in the social/behavioural domains but which are not identifiable with the methodology and philosophy of natural sciences. These considerations will figure in the ensuing discussion of internal criticisms of the scientific framework which will be presented in the next chapter.

Bearing this context in mind, then, the purpose of this chapter is to assess external arguments raised against the naturalistic approach. On closer inspection it will become evident that these arguments are based on faulty premises or errors in reasoning, or that they suffer from a combination of these two defects. The refutation of the basic arguments ranged against the possibility of a naturalistic social/behavioural science will, conjointly, furnish us with a "possibility proof" for the existence of such a thing. This is not so exciting as an indisputably compelling instance of theorizing based on this approach, or even as a full description of how such a piece of discourse would look. However, in the absence of any such example or any such detailed specification a "proof" of this kind is interesting.

A possibility proof in this informal sense should not be confused with the notion of an existence proof in formal languages, however. An existence proof establishes that something with certain defined characteristics must exist, even if no example has been found. A possibility proof is not so powerful; in effect it says that "it has not been established that a certain something cannot exist." There remains the logical possibility that a compelling argument

against the existence of the thing in question may arise -- though we may be in a position to judge that such a possibility is slim -- and there remains also the logical possibility of a successful general sceptical argument amounting to the conclusion that no method of inquiry can succeed.

It should be understood that an existence proof in the strict sense is simply too much to ask for in the present case. In a formal existence proof the statement that a certain construction must exist within a formal language, comprising analytical truths, is an analytical proposition itself. The statements (generalizations, laws) of a true empirical theory, however, are contingent, and the proposition that such true statements can be formulated must, by extension, be contingent also.

Chapter 3 will address certain issues that suggest defects in the scientific model or paradigm as it is currently constituted in our field. Essentially, the problems are as follows: (1) there is a lack of understanding of the characteristics of theory, and (2) educational technology research utilizes a mode of theory testing that does not support "strong inference" (Platt, 1964) ("crucial" experiments which eliminate one of two competing hypotheses while corroborating the other), that is not sufficiently strong as a method of winnowing out false hypotheses, and that simultaneously discourages the formulation of precise theories that might be tested more rigorously. The importance of this second discussion is that it suggests some reasons why the naturalistic approach has not produced any conspicuous successes, or any significant cumulative development of knowledge, to date, without invoking any of the anti-naturalist critiques that I will discuss presently. It provides a rationale, through "internal" criticisms, for the limited success (failure, to some minds) of the scientific approach without forcing the conclusion that naturalistic methods are doomed to failure.

### *The Arguments*

The arguments presented by the critics of the naturalistic approach break down into a number of categories. They turn on the alleged impossibility of separating fact from value in the domain of human behaviour, on the supposedly ineluctably context-dependent character of behaviour, on the nondeterministic nature of behavioural and social phenomena, and on certain

misconceptions regarding the manner in which theories are conceived and constructed in science. A more complete catalogue comprises the following:

- (1) Arguments proceeding from the alleged "value-ladenness" of social or "soft" science research (Howe, 1985; Guba, 1982).
- (2) Arguments from a "principle of uncertainty" or the alleged inoperativeness of causality in social and behavioural phenomena (Tranel, 1981; Guba & Lincoln, 1982).
- (3) Arguments proceeding from the notion of generalization. The conclusions drawn by Cronbach and Snow from their ATI research will figure here (Cronbach, 1975, 1985, Snow, 1973, 1977), along with arguments concerning psychological determinism presented by Lewis White-Beck (1974).
- (4) Arguments addressing the manner in which scientific theories are derived and formulated (Guba & Lincoln, 1982).
- (5) An argument based on the premises that education is an applied field, concerned primarily with change, and that the scientific approach is limited to answering theoretical questions. This argument concludes that a naturalistic approach is irrelevant to educational issues.
- (6) Arguments based on the alleged complexity and variability of social-behavioural phenomena.
- (7) Arguments from the alleged multiplicity of social "realities".
- (8) Arguments concerning the grounding of scientific theories, or the relationship between scientific theories and their associated data (Guba & Lincoln, 1982; Guba, 1979).
- (9) Arguments based on alleged limitations of scientific or hypothetico-deductive reasoning when compared with other modes (e.g., so-called hermeneutic reasoning).
- (10) Arguments which question the value of theory based on the notion that theories are radically underdetermined by available evidence and that they are, therefore, themselves arbitrary cultural constructs or artifacts.

These objections will now be considered, in turn, in the order in which they appear above, subject to the following exception: I shall defer consideration of the various lines of reasoning

falling under category (10) until Chapter 5, where they will be rejected following the presentation of extended arguments.

Obviously, these different categories of argument, (1) through (10), are not entirely exclusive. In particular, (3), (6) and (7) all relate to the alleged problem of complexity in some manner.

### *1. The "Value-Ladeness" of Social Science Discourse*

(1.1). The basic argument against the naturalistic approach from value-ladeness asserts that in the social-behavioural domain there are no value neutral facts to be discovered, while maintaining that the existence of such neutral descriptive facts is a precondition of the applicability of naturalistic methods. Perhaps one of the most concise and transparent versions of this argument is the one to be found in Howe's "Two Dogmas of Educational Research" (1985), where one of the alleged dogmas referred to in the title is the fact-value distinction. In a heady fit of inference, Howe saddles the reader with the following passage of purported deductive reasoning:

The positivist construal of the fact-value distinction is merely a corollary of the more general observation-theory distinction. If the positivistic attempt to ground all knowledge in some sort of atheoretical reality is untenable . . . then the justification for the rigid distinction between facts and values is equally untenable. (p. 11)

The problems here are manifold. First, it is simply wrong to say that the fact-value distinction follows from the observation-theory distinction. The more general distinction is actually the fact-value dichotomy, with the factual side subdivided into the theoretical and the observational. (The theoretical-observational dichotomy should not be construed as entailed by the fact-value distinction, however. The arguments supporting the two bifurcations are quite independent of one another.) So the argument is definitely not sound.

And neither is it valid, as it stands, for after stating that the fact-value distinction follows from the observational-theoretical distinction, Howe proceeds to say that the collapse of the latter entails that the former is not tenable. This is a glaring example of the error in reasoning called the fallacy of denying the antecedent. In an argument of the form "If P then Q",

ascertaining the falsity of Q is sufficient to refute P (deductive arguments are said to "retransmit" falsehood from conclusion to premises). But the falsity of P sanctions no conclusions whatsoever regarding the truth-value of Q.

Moreover, it is not correct to say that the observation-theory distinction is clearly untenable. Part of the logical positivists' program was the attempt to establish a firm foundation for all knowledge. This foundation was to take the guise of a peculiar variety of pure (noninferential) observation statement (sometimes called "protocols" or "sense data reports") which would be incorrigible or immune from error. This attempt failed and its demise helped seal the fate of "foundational" epistemology in general. However, less extreme (nonfoundational) forms of empiricism have been formulated, since the foundering of logical positivism, that do not require basic observation statements with the property of incorrigibility. And, within these more recent epistemologies, the observation-theory distinction has generally been upheld.<sup>1</sup>

Indeed, the observational-theoretical distinction, which was held in disrepute during the period in which Hanson (1958) and Kuhn (1964), its chief critics, acquired a large following, has since become something of a given again. The arguments against the distinction have not stood the test of time. The principal argument, which proceeds from a continuum of cases that extend between detection by direct observation and by inference, establishes at best that the distinction is a somewhat vague one. However, as Van Frassen (1980, p. 16) and others have pointed out, most of our predicates are vague, and providing that there are clear instances and counter-instances of a predicate it may be gainfully employed.

(1.2). Howe does offer some additional arguments against the fact-value dichotomy that are independent of his muddled reasoning from consideration of the observation-theory distinction, and these must also be weighed. He first issues the caution that it is impossible to prevent anyone using any term evaluatively. Thus, he concludes, "the very concepts social researchers employ are evaluative of human behaviour" (1985, p.12), and so the goal of an empirical, value neutral, science is incoherent. He goes on to concretize the argument:

consider the concept of intelligence and whether it is possible to strictly separate truth, facts and values. If research on intelligence involves solely the "goal of truth", it should



be possible to divest intelligence of evaluative meaning. is this possible? (1985, p. 12)

Once again, the line of thinking is hardly rigorous. For science does not require that we have no evaluative or emotive connotations associated with terms, but only that we can distinguish or isolate the descriptive meaning attached to them. We can, indeed, make this differentiation, just as surely as I can separate the question whether something is chocolate ice cream from the question whether I like chocolate ice cream or, perhaps, whether I think the practice of eating chocolate ice cream is unhealthy or even decadent.

Indeed, the distinction between fact and value is a fundamental characteristic of moral discourse, and one which has severely exercised the field of ethics. We frequently observe that participants in moral discussions may agree entirely as to the relevant facts, but still diverge in their moral judgements. It is a difficult task to develop a theory of ethics which is able to account for this aspect, and yet secure the notion that value judgements are somehow rational. (cf. Hare's *Freedom and Reason* (1963), for an extended discussion of this dilemma). An analogous problem arises in conjunction with other kinds of value judgements, such as the aesthetic.

But, again, all that science requires is that the terms addressed have a descriptive meaning and that this meaning can be separated from any of the various (possibly conflicting) evaluative and attitudinal connotations we may attach to them. This distinction which I have evoked between the descriptive import of a term and its emotive meaning is not a novel one. It was first suggested by the Cambridge literary critic I. A. Richards in his book published in the 1920's entitled *The Meaning of Meaning*.

It may be concluded, then, that Howe fails to undermine the viability of naturalism on the basis of either of his two attempts to show that the fact-value distinction collapses in the domain of social/behavioural phenomena: the one proceeding from the premise that the observational-theoretical distinction (which he wrongly identifies with the logical positivists' attempts to ground knowledge in incorrigible reports of the contents of sensation) is untenable; the other from the assumption value-free scientific discourse necessitates descriptive terms that can never acquire any evaluative or emotive connotations in general use.

## 2. *The Argument From the Principle of Uncertainty*

Tranel (1981) presents an argument against scientific method in social research that is possibly more inane than Howe's concerning value-ladenness, though it, too, has achieved some currency. (Guba and Lincoln (1982, p. 234), for example, indicate that they also subscribe to Tranel's objection.) The argument appeals to Heisenberg's uncertainty principle, one of the central elements in that branch of physics which deals with reality at the sub-atomic level Quantum Theory. The uncertainty principle addresses what happens when a measurement is taken of a system of sub-atomic particles. The act of observation exerts a photon pressure on the system observed, thus altering it. As a consequence it turns out that the actual behaviour of the observed system cannot be known with certainty; the best that we can do is ascribe a certain probability distribution. Tranel claims that the principle undermines the doctrine of cause and effect and that this has profound implications for the possibility of a social science. He writes: "If this (*viz* the failure of strict determinism) is so in the world of physics, it takes no great effort of the imagination to see how much truer it would be in the world of human uniqueness" (p. 426). In Tranel's view, this points to the incoherence of the very notion of a social science. He goes on to conclude that "If one can no longer speak of certainty and predictability and measurement in the areas of the physical sciences, then it is hardly credible that these notions would have any meaning in the social context" (p. 428).

There are myriad difficulties here. To begin with, Tranel has misconstrued the significance of the uncertainty principle. The first point that needs to be made is that Quantum Theory simply does not establish that reality, at the sub-atomic level, is inherently stochastic -- that the notion of cause and effect is inoperative. This is one possible inference, but what is entailed is only that measurement at that level is stochastic. The popular notion that Quantum Theory refutes the notion of cause and effect is due to the fact that the most widely held construal of the meaning of the formalism of Quantum Theory, the Copenhagen interpretation, suggests this. According to the formalism of Quantum Mechanics, it is impossible to assign *both* a determinate position and velocity to a fundamental particle. The Copenhagen interpretation of the formalism asserts that particles simply do not possess both properties at the

same moment. Other semantics for the formalism are possible, however. An alternative interpretation would be that particles possess both properties but that epistemic limitations (fundamental limitations on the extensibility of our perceptual and conceptual apparatus, imposed by the structure of physical reality) exclude our measuring both simultaneously.

The second problem with Tranel's argument is that even if it were known that reality at the sub-atomic level is stochastic, this would not imply that determinism does not apply in the case of social phenomena. The inference would go through only if it were possible to deduce the laws concerning social and behavioural phenomena from the fundamental laws of physics, and this is widely regarded as an unlikely eventuality.<sup>2</sup>

Tranel also oversteps the bounds of sense when he suggests that "one can no longer speak of certainty and predictability and measurement in the areas of the physical sciences" (p. 428). True, at one level of discourse "certainty" no longer applies, but "predictability" and "measurement" are still useful concepts – even if they have become stochastic in nature, when previously they were not. The point made by Quantum Theory is that the measurement or prediction of a single particle event suffers irrevocably from indeterminacy. However, the concepts of exactitude and determinacy still survive the transition from classical physics. So long as we contemplate a sufficiently large system, then Quantum Theory describes the world in terms of laws which express extraordinarily precise probabilities.

Thus, what Tranel completely fails to appreciate, apparently, is that laws may be of two forms: strictly deterministic or stochastic. The laws governing human behaviour may be inherently stochastic – but still laws for all that. The peculiarities of Quantum Theory do not have any obvious implications for the possibility of a social science. If anything, they make the notion of such a thing more plausible by demonstrating the possibility of a science whose fundamental laws are not strictly deterministic in form. Moreover, it is not necessary to contemplate post-classical physics in order to be struck by this insight. Classical physics was itself able to develop a very precise and powerful theory of the behaviour of gases based on probabilities, in the form of the kinetic theory of gases.

### 3. *The Impossibility of Formulating Reliable Generalizations in Social Scientific Inquiry*

(3.1). A number of arguments that have been adduced by the critics turn on the notion of generalization. Guba and Lincoln (1982), for example, maintain that reliable generalizations cannot be formulated in the social domain. In support of this claim they offer the following reflections:

Even in the hard sciences, however, there is a real question whether generalizations can be made that will be true "forever". Cronbach (1978) poses an interesting metaphor, that of the decay of radioactive materials, to make the counterpoint. Generalizations, he asserts, like radioactive substances, decay or have half-lives. He gives numerous examples from both the hard and the social behavioural sciences to make this point -- for example, the failure of DDT to control pests as genetic transformations make them more resistant to the insecticide; the shifting of stars in their courses so as to render star maps obsolete. (p. 240)

This passage is not very convincing; it seems to betray a lack of understanding of the nature of science. The goal of science is not so much to formulate reliable generalizations, (which may be only correlational) as it is to formulate laws. Laws differ from mere generalizations in that they are nomothetic. They support counterfactuals; whether deterministic or stochastic in form, they are characterized by their necessity. This reflects the circumstance that (unlike mere generalizations) they are derived within, or are part of, an integrated theoretical framework that is explanatory as well as predictive and descriptive. Mere generalizations stand essentially alone. They are descriptive and predictive, but they do not explain themselves

A good example to illustrate this difference is to be found in the progression, in classical physics, from Boyle's Law to the molecular or kinetic theory of gases. Boyles' Law simply relates the three variables pressure, temperature and volume according to the formula,  $pV = K$ , which asserts that at a constant temperature the pressure of a fixed mass of gas is inversely proportional to its volume (K is a constant which depends on the temperature and the gas) Boyle's law posits an invariant relationship among these variables, but it does not account for that relationship. The kinetic theory of gases, which succeeded Boyles's law has, as a

consequence, the same relationship. However, the theory puts forward a model of gases as comprised of clouds of molecules with certain properties which behave and interact in a certain way. This model provides an explanation of the conformance of the observable variables to Boyle's Law.

The generalizations Cronbach refers to in the passage cited by Guba and Lincoln do not constitute science, although they may have been formulated with the help of scientific knowledge. They are really technological artifacts. For example, star maps for a given point in time may be formulated on the basis of knowledge derived from that branch of physics called kinematics. But the theory itself actually explains why star maps become obsolete over time, because it comprises a framework of propositions that explains the behaviour of celestial bodies.

Moreover, there is no need, in order to justify the program of science, to assert that the laws which are formulated must be "true forever". Truth plays, in a certain sense, a methodological or regulative role in science. Science aims at truth, but approaches it asymptotically. The laws of science are tentative, conjectural. They are open to revision, not incorrigible, but this does not call into question the program of science. The essential characteristic of the methodology of experimental science in the natural sciences is that it is "self-correcting". False theories are refuted by the results of experimentation. Thus, while no number of experimental confirmations can guarantee the truth of a theory -- a false theory may entail many true hypothesis -- a single refutation is decisive and occasions new conjectures. To use Popper's phrase, it is this pattern of "conjecture and refutation" which enables progress (Popper, 1969).

(3.2). While the metaphors presented in the passage from Cronbach's 1975 essay, "Beyond the Two Disciplines of Scientific Psychology," chosen by Guba and Lincoln are unfortunate ones, there is still an argument raised by Cronbach which stands apart from those metaphors that must also be weighed. In that essay, Cronbach reflected back upon some two decades of research in the area of aptitude by treatment interactions (henceforth, ATI). What is striking about this body of research is that it indicates that generalizations about ATIs are not stable; ATIs vary from one context to another and, perhaps more disturbingly, also change over

time within the same context. This lack of stability in ATI results lead Cronbach and, to a lesser extent, Snow, his colleague and another major figure in the field of ATI research, to despair of the possibility of deep theoretical understanding in the behavioural arena. What was left for science, according to Cronbach, were merely the tasks of (a) "assessing local events accurately. to improve short run control" and (b) providing fruitful explanatory concepts that will "help people to use their heads" (p. 127).

Notice that this conclusion rests upon a certain conception of how deep explanatory theories are forged. Implicit in the reasoning is the idea that one must start with a certain critical mass of reliable lower level empirical generalizations from which may arise, perhaps by a straightforward inductive process, a subsumptive, systematizing, explanatory theory. Yet arguably this is not how most theory arises. The theory of evolution, while far from a perfect example of scientific theorizing, illustrates how one may arrive at a deep explanatory framework in the absence of a fund of empirical generalizations which are unaltered through time. Evolution deals with an enormously complex phenomena exhibiting a multitude of interactions, and these interactions vary over time. Yet it provides a theoretical account of the basic mechanisms which underlie the structure of these interactions. Thus, the pessimism of Cronbach (1975, 1985), Snow (1973, 1977), and others on this point may emanate largely from a naive conception on their part of how theories are created.

(3.3). In an earlier publication Guba (1979) presented another argument against naturalistic social science which turns on the notion of generalization.

While the concept of generalizability is appealing, it is also misleading. For if a generalization is a context-free statement, what can that mean for human behaviour which is always, inevitably, context-mediated . . . One can easily conclude that generalizations which are intended to be context free will have little that is useful to say about human behaviour. (p. 271)

Human behaviour is, indeed, determined by context; we can hardly take issue with this point. But so, too, are physical phenomena. In the latter case, statements of initial conditions establish whether the requisite conditions are present in a certain situation for a particular law to

obtain, and provide the parameters on the basis of which predicted values for experimental variables will be determined. The question, then, is not whether human behaviour is context-mediated. It apparently is, just as the behaviour exhibited by physical systems. The question is whether or not this mediation has any underlying structure, whether human behaviour is "lawful" in nature. This particular argument against the possibility of laws of social or behavioural phenomena thus falls because it simply rests on a conflation of the notions of generalization and statements of initial conditions.

(3.4). A more interesting argument concerning the impossibility of formulating behavioural laws which also turns on the notion of generalization comes to us by way of an attempt at refuting psychological determinism formulated by Lewis White-Beck (1975). Beck develops his argument in the context of considering the relationship or the dialectic between what he calls "actors" and "spectators". An agent is a spectator when she assumes the role of observing and trying to explain human behaviour. The agent who is observed is cast in the role of actor *vis-a-vis* the spectator. Beck comes to the conclusion that there is no possibility that a spectator may know a law to which agents must necessarily conform (pp. 127-128).

The argument goes thus: Suppose a spectator had established inductively the truth of a law, L, stating that "Actors in state S under condition C do A," via the behaviour of actors in state S under condition C. The actor who may suppose that he is determined because the spectator has established the law L is in fact deceived. Such an actor correctly believes that the spectator knows L. This very circumstance effectively differentiates him from the actors who constituted the basis for the formulation of the law, and so excludes him from L. The actor in question is not in state S, hence L does not apply.

Granted, the spectator may make a further induction based on a sampling of actors who are familiar with L, yielding a new law, L<sup>1</sup>: "Actors in state S<sup>1</sup> (S modified by knowledge of L) under condition C do A<sup>1</sup>." And this process of deriving a series of successive inductions can be continued as often as required. However, so long as the agent knows the last law formulated, he can exempt himself from it. It should be emphasized that this is an argument against psychological determinism. It leaves untouched the hoary issue of metaphysical determinism, an

issue which Beck says is undecidable (p. 129).

The argument obviously has interesting implications for the possibility of a social science. Is a science of human behaviour possible only on the condition that the laws governing behaviour are known exclusively to an elite group, who are themselves outside those laws?

The answer to this question may depend, ultimately, on how we construe the domain of a social science. Are the objects of a social science to be regarded appropriately merely as actors; or must they be conceived of both as actors and as real or potential spectators? We might distinguish between a first-order theory which addresses itself to an observed system, and a second-order theory which has as its object a -- potentially at least -- (self-) observing system. If the correct model for a theory in the social sciences is in the latter category, then Beck's argument against psychological determinism fails; it is irrelevant. And the possibility of a general social science survives.

More precisely, what Beck fails to take into account is that there is no reason in principle why a social or behavioural theory could not contain, among its postulates, hypotheses that (a) assert that theoretical knowledge of the determinants of social behaviour affects social behaviour, and (b) represent conjectures concerning the nature and effects of this interaction. Such hypotheses could potentially be integrated within a social scientific theory and subject to corroboration or falsification through the usual avenues of theory testing. In short, they would constitute empirical hypothesis on all fours with other empirical conjectures.

In brief, then, the standard objections raised against the possibility of forming reliable generalizations in social science areas turn on certain misconceptions: (1) the confusion of statements of initial conditions with empirical generalizations, (2) the erroneous view of theory-building which maintains that theories are constructed inductively from a stock or succession of stable, lower level empirical generalizations and, finally, (3) the presumptive belief that a social/behavioural theory could not be reflexive and could not include hypotheses about the nature of this reflexivity.

#### *4. The Sources of Scientific Knowledge*

Let us turn now to an argument against the naturalistic approach which attacks the



fashion in which theories are (allegedly) derived and formulated within the methodological framework of science. Guba appears to be a strong proponent of this style of refutation. Guba and Lincoln (1982) aver that: "Rationalists confine the types of knowledge admissible in any inquiry to propositional knowledge (Polyani, 1966), that is, knowledge that can be cast into language forms (i.e., sentences) . . . Naturalists [i.e., qualitative researchers] also build on tacit knowledge -- intuition, apprehension" (p. 245). Here, the term "rationalist" is used to designate one who believes in the naturalistic mode of inquiry.<sup>3</sup>

If the authors are, as it appears, trying to suggest that intuition, apprehension, and insight have no role in the development of science, then they are quite wrong. These phenomena play a decisive role in theory construction. The formulation and development of a theory is a creative act in the full sense of that term. But, by and large, these aspects of science are left out of discussions of methodology, for the simple and legitimate reason that nothing very precise can be said about them. While there exists a logic of the justification of theories, there does not exist a logic of the discovery of theories. The invention of a theory is a creative act, a psychological event, and little is known of the determinants of this process. Natural scientists and naturalistic social science researchers do discuss their intuitions -- in correspondence with colleagues, in seminars, in their biographical writing. They simply don't include such details in the published reports of their research. This is a good thing, in general, since the calibre of a theory (its empirical adequacy, its predictive and explanatory power, its scope and elegance, etc.) is not tied in any obvious way to the personal psychological history of the person or persons who created it.

It is disturbing to note, in connection with this particular attack of the scientific approach, that the critics actually proceed to retract it in a footnote of their paper (p. 251, note 6). There, they admit that intuition is crucial to naturalistic science but that the "reconstructed logic" of this approach precludes its public acknowledgement. They then go on to comment that: "The rationalist's unforgivable sin is to own up to humanness." It is serious enough that they miss the point of this reconstructed logic, namely, to render the justification and assessment of hypotheses as transparent as possible. What is more unforgivable is that they have, in the

context of the body of their paper, suggested that a genuine difference exists between the bases of theory formulation in the naturalistic and qualitative approaches and that this difference confers a preferential status on the latter, while in effect withdrawing the argument in a footnote. This seems dishonest. Perhaps what is most perturbing, though, is the charge that the naturalistic approach should be associated with a certain lack of humanity. It is just one of many indications that the attacks on naturalism are often rooted not so much in reasoned deliberation and epistemological argumentation as they are in ideological commitments. It is evidence of bad faith.<sup>4</sup>

Finally, it seems that the notion of tacit knowledge, or knowledge which cannot be cast into language, which is crucial to this argument, is a contradiction in terms. From a psychological viewpoint we know that some of our knowledge and skill can become so highly internalized and automated ("compiled", to use the term favoured by cognitive psychology) that we cannot, without serious effort and perhaps also external assistance, render it explicit. But knowledge is discursive and essentially public, and compiled "knowledge" can, in principle, be decompiled or reconstituted in propositional form. If something is ineluctably ineffable then it cannot represent knowledge. The only alternative to this view is a retreat into mysticism and obscurantism. But, then, whereof one cannot speak, thereof one should be silent.

##### *5. The Irrelevance of Theoretical Inquiry to Practical Educational Problems*

E. H. Carr (1983) has argued that educational research cannot be discussed within the usual framework of methodological debate in social science. He says that this debate, which contrasts two conceptions of educational research (naturalistic versus qualitative), actually begs the question whether educational research can be conceptualized as falling within the realm of social scientific inquiry: "It is an assumption that the philosophical basis of educational research can be understood in terms of this debate" (p. 36). Carr goes on to stipulate that education is not a theoretical activity but a practical one. The problems which it addresses are always practical and "as such, cannot be resolved by the discovery of new knowledge." Theoretical activities are defined, in this argument, as those which aim to discover something, while practical activities are those intended to bring about change. The argument is addressed to education,

but educational technology is at least as much a domain of practical problems as the broader field within which it is contained. We are, perhaps more so than educationists, fundamentally concerned with the design and implementation of educational or instructional systems and artifacts.

There is obviously a serious equivocation lurking in this line of reasoning. On the one hand, one can define discourse as theoretical or non-theoretical. Commonsensical and practical observations or recommendations fall within the latter category. On the other, one can define problems as either "theoretical" or "practical" to the extent that they represent either questions of knowledge or questions requiring that a decision or action be taken. Let us designate these two uses of the term "theoretical" as theoretical-1 and theoretical-2, respectively. With this equivocation exposed, Carr's reasoning can be revealed as the *non sequitur* which it is: "Social scientific discourse is theoretical-1. However, educational problems are not theoretical-2 problems. Therefore, social scientific discourse is irrelevant to education."

What Carr has obviously failed to establish, and what is essential to reaching his conclusion, is that theoretical discourse is necessarily irrelevant to practical issues. And there is clearly no good reason to presume that a convincing argument could be built to support that contention. Technology, after all, is often based on theoretical knowledge, and applied science certainly is. The question Carr should really be posing is whether a technology of education, or an applied science (of sociology, psychology, or whatever) is possible. For example, a detailed theory of learning *might* provide the basis for an effective technology or applied theory of instruction. But Carr does not pose these meaningful questions. Instead, he tries to extract further mileage from the distinction between practical and theoretical problems. For instance, he proceeds to note that theoretical problems are always determined by the theoretical background against which they arise. So, for example, "psychological problems about learning (which are of interest to the educational practitioner) are not determined by the practical problems experienced by learners" (p. 37).

In point of fact, however, one criterion which can come into play in deciding which theoretical problems will receive priority, whether we are talking about the natural or the social

sciences, is the question of which ones are most closely tied to what are perceived to be significant human problems. Moreover, as previously mentioned, it is also sometimes the case that theoretical questions are formulated in direct response to technological progress. Thermodynamics received a certain impetus, in terms of the conceptualization of thermodynamic processes and the structuring of theoretical questions, from the development of the steam engine. And, in an example closer to home, much of the theoretical work carried out in conjunction with tutorial, instructional and learning processes in the last decade has originated in the "practical" contexts of developing intelligent tutoring systems and of refining instructional programs and strategies that are intended to impart certain kinds of expertise and problem-solving abilities (cf. Wenger, 1987; Glaser, 1990). Such programs and technologies have, as we shall argue again later, provided the best opportunities and testbeds for developing and refining theories.

Carr's statement that psychological problems in learning theory are never influenced by the actual difficulties experienced by learners is thus far too strong. It should also be remarked that a requirement of a truly adequate theory of learning is that it should perhaps account for these difficulties, that it should be able to explain their genesis in terms of underlying mechanisms of learning. In other words, a complete theory of learning might well provide the basis for understanding learning difficulties, and for formulating a statement of the conditions that are crucial to successful learning.

In the last analysis, then, Carr's argument is extraordinarily simplistic and it fails to support his conclusion. Practical problems can certainly be distinguished from theoretical ones, according to the definitions he supplies. But decisions or actions are taken, hopefully, in light of knowledge, and such knowledge might be comprised, at least in part, of theoretical insights. Surely they will be taken, where possible, with due consideration paid to relevant generalizations, and theoretical knowledge is one important source of such information.

#### *6. The Intractable Complexity of Social Phenomena*

It has become commonplace to say that naturalism cannot be fruitful in the social/behavioural domain because of the sheer complexity of the phenomena. The argument

generally asserts that there are too many variables in the social context, that there are too many interactions among these variables, and that it is too difficult to isolate the phenomenon under investigation from other, confounding, causal influences, for the experimental approach -- which is predicated on the possibility of isolating a rather restricted number of explanatory variables, and manipulating these to determine their roles -- to work.

In some instances the premise that human behaviour is context dependent is adjoined to further the case for intractable complexity. I have already shown (section (3.2), above) that this last point rests on a conflation of generalizations with statements of initial conditions. Whatever exact combination of premises is brought to bear, the argument from complexity eventuates in the claim either that it is impossible to isolate the variables of interest (or to know that we have isolated suitably similar instances of the same variables) or, less radically, that social scientific generalizations can only ever be descriptive, encapsulating past observations (the changing context and patterns of interaction precluding prediction or nomothetic law statements).

Such arguments, advanced both by proponents of qualitative inquiry and by systems thinkers, are ultimately not very convincing. To begin with, it should be noted that the appearance of intractable complexity may simply reflect the lack of an integrative, simplifying theory of the phenomena. Without a theory any variability is potentially significant. It is only as theoretical conjectures are advanced that we exclude some of this potential. So to say that a system is too complex to investigate by experimental methodology is possibly only to confess that we have no theory, no conjectures, in hand. There is an indeterminate degree of variation in nature, no less so than in social systems; the difference is that we have developed simplifying, integrative theories that show how much of this variability can be explained with reference to a small number of theoretical constructs and a limited number of principles. One cannot rule out *a priori* the possibility of achieving the same in the social/behavioural sciences. We have also seen (section (3.1), above) the possibility that patterns of interaction among relevant variables are transformed over time is not fatal to the idea of naturalistic theorizing. Moreover, so far as the question of whether apparently diverse manifestations are to count as exhibiting the effects of the same variables and instantiating the same generalizations under different initial conditions,

or whether they really represent different phenomena, is concerned this can only be decided through the process of experimentation and observation and the testing of conjectural generalizations (Popper, 1969).

### *7. The Multiplicity of Social Realities*

There is one final argument which is not unrelated to the argument from complexity that should be noted, namely, the argument from the multiplicity of social "realities." Guba and Lincoln (1982) maintain that the scientific approach cannot succeed because social reality is a fragmented phenomenon. They claim there is no single objective, independent social reality, but that rather there exists an indeterminate number of subjective, socially constructed realities "the realities are multiple (as many constructions as there are people), it is futile to expect convergence" (p. 239).

There are difficulties connected with this view which become more transparent when we ask what logical consequences can be drawn. If we were to take this view, extreme as it is, seriously, how could we account for the possibility of communication? Guba and Lincoln present a view of the world in which we are each trapped within the confines of our own private reality. The position is extraordinarily solipsistic. It calls into question not only the possibility of a science of social phenomena, but also the purposes of qualitative inquiry. To begin with, by Guba's hypothesis, the outcome of the latter enterprise is a description of the individual case. What we are now told suggests that there must be as many different and potentially irreconcilable descriptions of the individual case as there are interested parties. This relativizes the notion of explanation and thereby undermines the notion of qualitative "theorizing"

Also, the very possibility of ever developing an interpretive understanding would seem to preclude this radical conclusion. Davidson (1963) argues for a "principle of charity" as a basis for understanding the behaviour of members of different cultures. According to this methodological precept, we should seek to maximize the agreement between what we believe and the beliefs we attribute to others. MacDonald and Pettit (1981), alternatively, have advocated a "principle of humanity" whereby we seek to minimize the degree of unintelligible disagreement between the beliefs we hold and those we attribute to informants from any culture

we are trying to understand (p. 29). This is a weaker principle than Davidson's insofar as it does not carry the implicit assumption that others think as we do. Rather, it assumes only that others are "behaviourally rational" (meaning that some of their behaviour at least constitutes actions which follow from beliefs and desires that serve to rationalize those actions) and "attitudinally rational" (indicating a certain disposition to change beliefs in light of the discovery of inconsistencies or lack of fit with experience).

Nonetheless, even this weaker precept carries with it the implication that there is some commonality among the beliefs and attitudinal systems of rational agents, even where the agents are the products of diverse cultural frameworks, and that this common ground, however narrow, must serve as the basis from which any attempt to develop an understanding of the behaviour of others must proceed.

Other difficulties intrude on the qualitative inquirer's own terms. Collin (1985) has argued that, even apart from the problems of relativization and methodological solipsism, qualitative or interpretive ("belief-desire") explanations of behaviour cannot provide a theoretical understanding of human behaviour because they are necessarily so variegated. Collin begins his discussion with an analysis of the notion of theoreticity. There is a sense of the term, which I will explore further in the second part of this chapter and which has already been introduced above, in which theoreticity is contrasted with the notion of observationality. In this sense, the term marks a distinctive epistemological category: factual knowledge is classified as either observational or as theoretical, essentially on the basis of whether it is given directly in perception or whether it must be inferred.. There is another sense of the term, however, which Collin utilizes in his assessment of interpretive social inquiry. In this sense, theoreticity refers to the "integrative, organizing power of knowledge systems, their power to compress knowledge into an easily surveyable form" (p. 60). Theoreticity, defined in this manner, is thus a property possessed in varying degrees by all systems of knowledge, and exemplified in the highest extreme by certain examples encountered in the physical sciences. The concept is closely allied to that of explanation: those systems of knowledge which rank higher in theoreticity also rank higher in explanatory power.<sup>5</sup> In the process of subsuming what, at first blush, might appear to be diverse

phenomena, under a single, smaller set of abstract principles, we effect more powerful explanations of these phenomena. Newton's concept of universal gravitational attraction is one of the most compelling illustrations of this relation.

There are actually two forms of integration which occur as knowledge systems slide up the scale of Collin's "theoreticity". The first is the integration of principles mentioned above; the second is a further reduction and consolidation of the ontological commitments they express. Both forms can be judged along two separate dimensions: generality and parsimony. Often an increase along one dimension will coincide with a movement along the second in the same direction. For example, Newton's theory postulated one reduced set of more general principles to account for both the motions of bodies on the surface of the earth and the orbits of the planets, phenomena which previously had been considered unrelated. At the same time, his notion of gravitational attraction, the basis for this reduced set of principles, also afforded an ontological reduction in science.

This positive relationship between generality and parsimony is not a necessary one, however. It is possible, for example, to have an increase in generality, through the addition of principles that are broader or more encompassing, coupled with a decrease in parsimony, if these new principles do not actually render the existing, narrower principles superfluous. But by and large, the two will tend to vary together; it is the discovery of more subsumptive concepts which allows for the possibility of managing with a reduced set of principles or rules. It is clear from this discussion that what Collin intends by the term "theoreticity" is closely allied with the concept of "simplicity". Collin's point about purposive or interpretive explanations of human behaviour as actions is that they are not amenable to the types of integration that lead to greater degrees of "theoreticity" and explanatory power. The problem lies in the sheer and irreducible variety of human beliefs and desires. To appreciate this point fully we must begin with Collin's premises concerning the completeness of rational or interpretive explanations of behaviour. Collin raises the question whether the everyday intentionalist language in which we explain our behaviour as actions (i.e., as behaviours imbued with meaning, as behaviours intended to achieve certain goals or satisfy certain desires, in a rational way given our beliefs), would permit



of theoretical generalization:

Does not the action language allow us to introduce a more general term so long as this term *subsumes* the relevant class of more concrete, everyday action descriptions?

Such subsumption would not impugn the validity of these descriptions; on the contrary, it would presuppose it. (1985, p. 214)

The answer Collin provides is negative. His response turns on the assumption that everyday accounts are already "maximally specified," a phrase by which he intends that they convey the full description under which the goal is desired. If this condition is satisfied, there is no room for further generalization; if not, then the intentionalistic explanation is inadequate as an everyday explanation:

A true, maximally specified everyday action description will not merely indicate the specific desirability properties of the action (or its upshot), but will cite its generic desirability properties as well. And if the specification is indeed maximal, no more universal term may be added to it. (p. 214)

Nor, he argues, is this conclusion undermined by the fact that we may be able to find, for purely classificatory purposes, a more general category under which a given action may be subsumed. Such a more general rubric will not necessarily refer to the desirability characteristics of an action. Consider the example Collin uses to illustrate this point. Suppose an agent has a desire for a glass of cold milk and for a new automobile. Now introduce the new concept of something which is either a glass of cold milk or a new automobile, and designate this by the term "mobilk". It would be wrong to say that the agent has a desire for mobilk, for whenever the notion of satiation applies with respect to a desire, then it is part of the logic of belief-desire explanations that attainment of the object of that desire leads ultimately to a decrease in the intensity with which further objects of the same kind are sought after. The agent's alleged desire for mobilks clearly does not meet this condition. Acquiring a glass of milk will not reduce the longing for a new car, or vice versa. Thus, the disjunctive class, mobilk, cannot be regarded as a genuine object of desire and so cannot appear legitimately in a rational account of the agent's actions.

It seems, then, that a strong case can be made that interpretive inquiry cannot lead to the formulation of highly integrated explanatory frameworks, featuring precise, reliable generalizations. This, of course, strengthens the objections to any program to supplant quantitative research with a qualitative approach.

#### 8. *The Superiority of "Grounded" Theory*

Guba (1979) also argues for the superiority of the "grounded theories" generated by qualitative methodologies over the "*a priori*" theories which evolve within the scientific approach. Grounded theories, he claims, are "theories derived from real world data and information" and are to be contrasted with "theory stemming from *a priori* assumptions in the form of the typical hypothetico-deductive theories of science" (p. 271). Elsewhere he continues this line of attack, asserting that "the naturalist (read "qualitative researcher") does not search for data that fits his or her theory but develops a theory to explain the data" (Guba & Lincoln, 1982, p. 235). Finally, he concludes that: "In all events, theory is more powerful when it arises from the data rather than being imposed on them" (Guba & Lincoln, 1982, p. 244).

There are two significant aspects to these passages. In the first place, it is inaccurate to say that scientific theories are *a priori*. It is true that sophisticated theories cannot be derived deductively from any finite set of data. A scientific theory goes beyond the data presented. It addresses an infinite number of cases. Moreover, the semantics of a powerful theory will characteristically include hypothetical constructs, terms which designate entities and processes that cannot actually be observed in the data, but which somehow explain at a deeper level the patterns which can be discerned there. Since the semantics of a theory will generally go beyond what can be directly observed in the data to which it refers, the theory cannot even be arrived at by a mere process of induction by enumeration, the *modus operandi* of lower level generalizations.<sup>6</sup>

But, for all that, theories are empirical in nature and to describe them as essentially *a priori* is to suggest, wrongly, that theories are either (a) created entirely by rational intuition, without any initial constraints imposed by what is observed, or (b) that they are analytical in nature and are based solely on definitions.<sup>7</sup> To suggest that theories are essentially *a priori* is to

misrepresent the processes of science and the subtle interplay between theoretical conjecture and its associated data. It may also be that Guba again confuses the context of the testing of hypotheses with the context of their discovery. The hypothetico-deductive model to which he alludes is really essentially an idealized model of how theories are confirmed or justified. It does not speak so much to the more inscrutable dialectic of the processes concerned in their elaboration.

In the second place, it is most interesting, as well, to read Guba's contention that theory generated by qualitative methodology must be more powerful than theory generated by the scientific approach. By the admission of many of its proponents, qualitative approaches do not generate theory at all. Guba himself, for example, tells us that: "The aim of inquiry is to develop an idiographic body of knowledge. This knowledge is best encapsulated in a series of 'working hypotheses' that describe the individual case" (Guba & Lincoln, 1982, p. 250). The knowledge uncovered by naturalistic methods is thus idiosyncratic, relevant only to a particular context. This is not "theory" at all, and it is certainly less powerful than the law statements and deep explanatory frameworks that comprise scientific theories and that are epitomized by the highly successful creations encountered in the hard sciences.

*The prospects for generalizable non-naturalistic theorizing.* Having raised the question of the feasibility of qualitative theorizing once again -- I have already presented Collin's objections, which were predicated on the sheer variety of intentions and values -- a few words concerning the general tenor of the philosophical literature on this point are apposite. In the philosophical literature, at least the portion in the analytical mould, so much attention has been focused on the question whether there can even be a distinctly noncausal form of explanation which would differentiate naturalistic and qualitative ("interpretive" or *Verstehen*) approaches that it has generally been assumed that establishing the independence of qualitative explanations of behaviour would credit the possibility of generalizable non-naturalistic theorizing. Yet there is ample reason to doubt this implication.

The locus of discussion on this question has been Davidson's work on the theory of action and a semantics for social science. (Davidson, 1963; 1964; 1967; 1971). Davidson

of variables controlling action, only the intervening variables are mentioned, whereas the independent variables which control action through the intermediation of these variables are left out. (pp. 337-338)

Collin's objections to construing purposive accounts as inherently causal then turn on the possibility of cases of deviant causation. To make his point he sketches a case where a character named Bill kills his uncle as a causal consequence of his desire to obtain this relative's money, together with a belief that he can acquire it, via inheritance, by killing him. Now consider two alternative scenarios. In scenario one, Bill arrives at his uncle's residence, finds the man and coolly and deliberately dispatches him with a revolver. In scenario two, we find Bill in such a state of tension that, on producing the gun, his nervous agitation causes his finger to jerk, squeezing the trigger and sending his uncle to the hereafter just as effectively as in the first scenario.

Observe that the purposive explanation posited for Bill's behaviour is equally tight, in causal terms, for both scenarios. In both instances, a causal link is established between the event which occurs and Bill's conative and cognitive states. And in both instances the same intentionalist states are implicated. Yet Collin argues that the two accounts do not have the same "force". And this difference in force, he maintains, establishes that there is indeed a form of purposive explanation that is distinct from straightforward causal explanations.

It will not do, Collin continues, to defend the thesis that there is no distinctive interpretive mode of explanation by trying to locate the differences in the explananda: to argue that in scenario two Bill "shoots" his uncle only in a tenuous or derivative sense "that does not presuppose that he fired the shot *in order* to dispatch his uncle," while the account is a genuine intentionalist one in the case of the first scenario. To try this avenue, he says, would clearly be to grant the interpretivist his case.

Another alternative, though, would be to say that there is a difference between the two accounts, and that the difference stems from the "rationalizing" effect of the first, but then deny that this rationalization is related to the explanatory function of the account. Collin has a rejoinder to this gambit also: rationalization cannot mean only the observation that an action is

introduces a technical notion, that of a "primary reason", in his account of interpretive explanations. A primary reason corresponds to a complex mental event comprising a pro-attitude towards actions of a certain kind together with a belief on the agent's part that he will perform an action of this type in the near future. This mental event causes certain bodily behaviour, and this behaviour constitutes the action, in virtue of its having the appropriate causal determinants as specified by reference to the primary reason. Thus, on Davidson's account interpretive or purposive explanations of actions involve a reference to mental states which in effect convey what we recognize as reasons, but these mental states are causally efficacious. Rational explanations are thereby revealed as a species of causal explanation, and another possible ground for distinguishing between naturalistic and nonnaturalistic explanation is jeopardized.

Collin has rejected this assimilation of rational or purposive explanation to the causal variety. His arguments do not seem to have much force, but it is instructive to examine them in so far as they highlight the difficulties that are encountered when one tries to defend this distinction.

He begins by allowing that interpretive or "belief-desire" explanations do invoke causal factors, but then proceeds to note that this alone is not sufficient to render them causal explanations as such. What is also required of causal explanations is that they derive their explanatory force from the attribution of causality:

When we refer to something as a causal explanation, we are not merely saying that it is an explanation and that it involves causation; we are saying that it is *explanation by causation*, that it is precisely the demonstration of a causal tie that delivers the explanatory power. (pp. 108-109)

He then goes on to characterize interpretive accounts and to argue for their distinctiveness. In explicating the notion of an interpretive explanation he begins by describing motivational states (akin to Davidson's primary reasons) as intervening variables which mediate the effect of behaviour on the agent's environment.

We may characterize interpretive accounts as explanations in which, out of the total set

rational given the agent's interests; if that were so, then the two accounts would be identical, since in both cases Bills's acts are equally appropriate given his purpose.

Rationalization must thus involve such appreciation of rational suitability, *plus* recognition of the fact that the action was caused in the appropriate manner by the cognitive and conative states of the agent. But given this interpretation, it is no longer obvious that rationalization has nothing to do with explanation. It no longer means the mere appreciation of certain abstract rational connections, but in addition the recognition of certain mental states as causally operative in the situation.

(p. 111)

But surely this is simply to return to the view that the explananda are different. And yet it is not clear that this plays into the interpretivist's hands. In one case (scenario two) what is required is simply a causal explanation (essentially a physiological one in this case), while in the other (scenario one), the causal explanation involves, as a direct causal determinant, a so-called "primary reason". But it is not so self-evident as Collin supposes that the intentionalist aspect, as edifying as it is, adds anything to the explanatory power of the account. Collin's response to this line of attack seems to be that the illocutionary force of the account is different, and that this difference has to be interpreted as a difference in explanatory power. Arguments based on the perception of gradations of illocutionary force are far from decisive, however. Collin has failed to locate the difference in the semantic properties of the two accounts and the feeling of greater force in the second account may be misleading; it may only be the sense of edification it inspires that influences our judgement so -- creating, as it were, an illusion of surplus explanatory force. It thus appears that the distinction between causal and interpretive explanations of behaviour, despite its intuitive appeal, is not an easy one to defend via claims of a differential role for attributions of causality in accounting for their explanatory power.

### *9. Hermeneutic versus Deductive Modes of Reasoning*

It is sometimes held that a different kind of understanding underlies qualitative inquiry than the naturalistic approach: "hermeneutic" appreciation rather than deductive explanation. The term hermeneutics originally referred to the scholarly interpretation of biblical texts. It has

since come to be used to refer to the interpretation of literary materials in general and, also, to designate the kind of circular or spiralling mode of thinking that appears to characterize much of interpretive social inquiry. Following the refinement and development of hermeneutical methods by Dilthey and Ast, the idea that hermeneutics is a distinct approach peculiarly suited to the study of meaningful materials -- i.e., anything that expresses an agent's beliefs, desires or values -- has become quite fashionable.

Deductive reasoning comes into play in the naturalistic sciences in the context of the logic of justification or more specifically in regards to the experimental testing of theories or general hypotheses. The idea is that one takes the hypothesis in question, along with all necessary statements of initial conditions and definitions, and generates a prediction by means of deductive logic. In other words, the prediction is a theorem of the hypothesis (or collection of hypotheses) concerned. If our experimental measurements or observations are incompatible with this logical consequence of the hypothesis, then essentially the theory has been falsified by the rule of inference, *modus tollens*. This general scheme is usually referred to as the hypothetico-deductive model of theory testing. This same scheme also implies a certain conception of explanation. A particular phenomenon has been "explained" if it can be shown that a proposition describing that phenomenon is derivable from some more general hypothesis (theory or law). Thus, the conception of explanation is one which is symmetric with respect to prediction.

The question whether hermeneutic understanding is really different from the hypothetico-deductive approach is an important one for two reasons: (1) if it is truly distinct then this raises the possibility of constructing an argument against the possibility of a naturalistic social science, based on the idea that understanding of human actions necessitates a hermeneutic approach which is incompatible with naturalism, and (2) while the various qualitative approaches diverge in certain respects (cf. Jacob (1988; 1987) and Anderson (1989) for an overview of the various qualitative traditions), there are two elements which seem common to all, whether they be founded on phenomenology, existentialism, or ethnomethodology (Smith, 1984, p. 381). One is an acceptance of epistemological idealism, and the other is this belief that the significance of

human actions must be grasped hermeneutically. Thus, the question of the nature and possible distinctiveness of the hermeneutic approach is central to understanding the differences between qualitative approaches and the naturalistic scientific approach.

The argument we must confront is one which asserts that, in the search for interpretive or qualitative understanding of human behaviour, this deductive or "covering law" model of explanation is inapplicable. Essentially the argument must run something like this: in qualitative inquiry, our explanations or "theories" must be "grounded"; that is to say, they must emerge from the data and must not be imposed upon it. On this account, patterns emerging as the data are collected will suggest tentative, working hypothesis that will guide, in turn, further selection of data and procedures. These working hypotheses will be in constant flux, and a circular pattern of contextual reasoning will emerge that is very much like what occurs in the analysis of literary works. The notion of holism plays a role here, also, for in textual exegesis there is a constant movement from the interpretation of the significance of the work in its entirety to the fathoming of specific parts of the text, and back again. The interpretation of the whole text influences the interpretation of the parts, but of course the meaning of the work in its entirety is also determined by the meaning of the elements. The balancing and reciprocal accommodation of interpretation at the global and local levels is generally referred to as the "the hermeneutic circle".

Follesdal (1979) has argued cogently that hermeneutics is simply the "hypothetico-deductive method applied to meaningful material (i.e., texts, works of art, actions, etc.)" (p. 320). He illustrates this by considering a concrete case of literary interpretation, examining five interpretations which have been advanced of the meaning of the stranger who appears twice in act five of Ibsen's *Peer Gynt*. In each of the five interpretations, it is clear that hypotheses are advanced, consequences are derived, and these consequences are compared to the text to assess their fit. In the process of deriving the consequences, auxiliary hypotheses, in the form of theories of literature or style, and additional information concerning e.g., the author himself, are brought to bear. The similarity to the process of advancing hypotheses and testing what they imply in natural science is quite clear. Moreover, in literary analysis there is a preference for



"unitary" interpretations, meaning those which establish connections among different parts of the text. This is analogous to the preference in science for "simpler" hypotheses -- by which we intend, at least in part, hypotheses which are able to account for a greater diversity of data.

I would agree that, on an intuitive level, the constant movement in hermeneutics between global hypotheses concerning the meaning of a work as a whole and more specific conjectures concerning local elements may seem distinctive as compared with hypothetico-deductive procedures in science. But any difference is really only apparent. In science, our theories influence our interpretations of the data, yet, at the same time, our observations influence our high-level theoretical conjectures. Indeed, this idea of a reciprocal influence operating between the interpretation of specific observations and the formulation of our global theories is the central tenet of post-positivist philosophy of science.

So the shift between interpretation at the two levels, global and specific, is widely perceived to exist in science as well. The difference is really that in textual exegesis, which proceeds more rapidly than the repeated development and testing of hypotheses by experimental means in science, the transitions from one level to the other are much more obvious or transparent. The comparison that should be made, though, is not the one which is typically posed, namely, what transpires in the analysis of a literary work against a single experiment in science, but rather that of hermeneutics versus a whole series of experiments during which many tests of predictions are concluded, and accommodations for recalcitrant observations must be made by way of either altering our interpretation of the evidence or modifying our conjectures. When the comparison is framed in this way, the appearance that the hermeneutic circle is something unique to reasoning in the humanities or interpretive social science dissolves.

Follesdal makes a number of additional points that are relevant here. First, he emphasizes that his analysis only supports the conclusion that hermeneutics is best described not as a separate species of understanding, but as an application of the hypothetico-deductive scheme to "meaningful" subject matter. It does not support the further conclusion that all forms of understanding are reducible to the hypothetico-deductive mode. Such a conclusion, he

urges, would be incompatible with the basic thrust of this approach, which is that all knowledge should be considered conjectural.

This leaves open the possibility that an argument might still be constructed to show that understanding of human activities requires a different *modus operandi*, one that is incompatible with the hypothetico-deductive approach. Based on the foregoing arguments, however, this special mode of understanding would have to be something other than the hermeneutical. One alternative which deserves mention would be a form of phenomenological understanding. But there are difficulties associated with this mode of thinking as well, at least so far as the goal of constructing a form of discourse which would provide general knowledge that is open to public verification is concerned. The basic idea behind phenomenology is that we condition our experience of the world through the categories and constructs inherent in the language by which we describe it. From the perspective of a "constructive" epistemology, such as deployed by Kant and his successors, this is inescapable, and the task of the philosopher is to clarify the basic set of categories which, in its role of providing a structure to sensation and a framework for thought, is a precondition of any form of experience characterized by some degree of order. Within the constraints of such an epistemology the notion of what an object is in and of itself (what Kant called the "noumenon"), apart from the application of the categories which serve to constitute it as part of an intelligible mode of experience, has no sense.

The goal of phenomenological method, however, is to progressively strip away the influences of the categories of language, through a series of "reductions", to arrive at a more direct form of intuition or apprehension of the objects of experience, unmediated by the categories and conditioning of language. The difficulty is that even if one allows that a phenomenological appreciation is possible, one must conclude that nothing can be said to describe or explicate that appreciation. The moment that one begins any form of explication the categories and structurings of language intrude themselves once more and the purpose of the exercise is defeated. A more radical form of criticism can also be advanced, focusing on the notion of understanding. It might be argued that the term "understanding" in the context of phenomenological method is misused, since what we mean by the term, in common usage, is

precisely the subsumption of objects under different conceptual categories conveyed in a language.

Follesdal also addresses certain arguments the upshot of which is that the hypothetico-deductive model cannot be applied to human actions. One of these turns on the claim that the hypothetico-deductive method presupposes that the researcher who applies the method does not affect the object or system which is investigated, as happens in social/behavioural inquiry. This is sometimes conjoined with the assertion that the hypothetico-deductive method cannot accommodate the possibility that the researcher himself is a part of the phenomenon he is investigating. The response to these objections can be expressed quite briefly: the hypothetico-deductive model can, in principle, accommodate such considerations quite easily. A social science might posit these assertions -- that the researcher influences the system he studies and that he is part of that which he studies -- conjoined with certain conjectures concerning, e.g., how these reciprocal influences between the act of theorizing and the nature of the object of theorizing operate. All these conjectures, taken together, would be assessed based on their mutual coherence and the tribunal of experience. Similarly, there is little sense to be found in the objection that the hypothetico-deductive method cannot be applied to self-reflection. There is no reason in principle why a hypothetico-deductive system cannot include sentences that refer to oneself and one's own activities, or even sentences which address the role of self-reflection in a hypothetico-deductive explanatory system (reflections on self reflections). Indeed, I relied upon this potential of hypothetico-deductive systems to include these sorts of conjectures in order to undermine Beck's sceptical arguments in section (3.4), above.

Thus, it seems there are no clear cut grounds for accepting that a difference exists between interpretive and causal accounts on either the basis of the mode of reasoning (hermeneutic versus hypothetico-deductive) employed or the extent of the role played by attributions of causation in terms of their explanatory power.

#### *Conclusion*

It seems that the standard battery of arguments levelled against the scientific approach by its critics are less than cogent. In some instances the premises of these arguments are either

not well-supported or are patently false; while in other cases we encounter logical errors in argumentation of a very basic and very transparent nature. All in all, the dismissals of the scientific approach which have been conveyed in the recent literature are rather too glib. Appeals to authority -- frequently sources in the philosophy of science that are now outdated and discredited such as Kuhn or Feyerabend, or individuals who have always existed on the periphery of the discipline such as Polyani -- are surprisingly numerous.

A favourite ploy is to describe some aspects of current scientific methodology as linked with logical positivism and then dismiss the whole of the scientific approach on that ground.<sup>6</sup> There is something to be said for this approach, since the views comprising logical positivism have been abandoned even by their original proponents as seriously flawed, and since some methodological aspects of the social/behavioural sciences do reflect parts of this older, unserviceable philosophy of science. But this does not show that the scientific approach cannot work. It only shows that it needs to be reformed, that methodology in the human sciences suffers from a certain inertia, having been defined and crystallized, in certain respects, some fifty years ago. Much water has passed under the bridge since the days of logical positivism. The philosophical analysis of the methodological principles of the hard sciences has brought us to a much finer understanding of the meaning of science and the foundations of its practice. At the same time, the methodological principles of the hard sciences have themselves undergone a certain evolution. They are always under question, always open to revision, as much so as the theories of science themselves. In fact, the refinement of methodology is essential to the advancement of theory. It is in this regard that the social sciences have failed, clinging implacably to some principles that are old and discredited. But it is a failure that may perhaps be remedied.

There is a point in this. We should indeed debate the merits of our methodology. We should adopt and maintain, at all times, a critical attitude. This is the only way we can refine our methodology and push our sciences forward. But the debate as it is currently being played out is not altogether a constructive thing. The tone is too emotional; there appears to be an ideological as much as an intellectual motivation at work behind the questioning (science is

somehow regarded as essentially inhuman or dehumanizing both with regards to its products and its practices). And, in general, the level of intellectual rigour exhibited is less than we should strive for in such a serious matter.

## CHAPTER 3

### Internal Criticisms of the Naturalistic Approach

In the previous chapter it was maintained that the arguments typically brought against the "scientific" model of inquiry in the social/behavioural sciences are not very strong ones, in the last analysis. In this current chapter I want to move on to suggest that while this is so there are yet conclusive reasons for believing that the scientific model as it is presently represented will not, cannot, yield truly substantial results. I also want to suggest that these reasons may not be widely acknowledged by researchers and practitioners in our field.

The difficulty is that the methodological framework generally identified as science in the field of education and psychology is not one which has ever been exploited successfully in any field of study. It is not, contrary to popular belief, a close approximation of the approach which has been so fabulously successful in the hard sciences, and which social scientists strive to emulate. It is based, rather, on an analysis of science (physics, in point of fact), its structure and cognitive status, which was put forward in the late 1920's. But this analysis was mistaken, gave rise to certain absurdities, and shortly was soundly repudiated within the field from which it originated. The position I refer to is called operationalism, a position elaborated by the Harvard experimental physicist Percy Bridgman and first presented in his book entitled *The Logic of Modern Physics* published in 1927. The position was soon adopted wholesale by the field of psychology in a very self-conscious way as a result of internal and external pressures to be "scientific". (A clear and concise statement of the position in the literature of psychology is provided by Bergman (1953).)

#### *The Defects of Operationism*

Operationism is, in brief, the view which asserts that every theoretical term in a theory, if it is to have a role in that theory, must be related to certain experimental operations and can be explicitly defined by these procedures. Now it can easily be shown that this view leads to absurdities. To take an example from physics, in the operationist view electrons and protons are construed as meter readings. We are then led irrevocably to assert statements of the following sort: (i) A hydrogen atom has one electron and one proton

=df A hydrogen atom has two meter readings.

Following the precepts of operationism we must now operationally define "hydrogen atom", so we obtain:

(ii) =df A meter reading has two meter readings.

One may well ask what it might mean to say that meter readings have other meter readings. What this illustrates, in short, is that it is absurd to define a concept which in itself has nothing to do with operations solely in terms of operations. One can certainly add to the understanding of a theoretical term, T, by all sorts of statements of the form:

(iii)  $(x)[Ox \rightarrow (Tx \rightarrow Rx)]$

Here we say that if we perform an operation O on a research unit x, then if x has the theoretical property T it will exhibit the observable property R. Such statements can expand our understanding of T, but do not in the end define it. We do not arrive by this path at an expression which can replace T without loss of meaning.

We could even have the stronger claim:

(iv)  $(x)[Ox \rightarrow (Tx \leftrightarrow Rx)]$

Here the response R is a necessary and a sufficient condition for x having the property T, but only if O is performed on x. But this still does not furnish a definition of T. To explicitly define T operationally is in fact to advance a stronger statement of the following form:

(v)  $(x)[Tx \leftrightarrow (Ox \rightarrow Rx)]$

This is clearly a semantic absurdity. Please note that:

(vi)  $(Ox \rightarrow Rx) =df (\sim Ox \vee Rx)$

That is, the conditional statement "If O is performed on x, then x exhibits R" is logically equivalent to the disjunctive statement "Either O is not performed on x, or x exhibits R." The disjunction is true whenever either of the disjuncts is true, so according to (v) we are logically compelled to ascribe the theoretical term T to any research unit to which we have not previously applied the operation R. Thus, for example, we must assert that a cube of jello is brittle on the condition that we haven't tested it for brittleness. Once again, the premises of operationalism have brought us to an absurd conclusion.

To summarize, there are two fatal errors involved in operationism. The first error is a

semantic one. This is the assumption that concepts which have nothing inherently to do with operations can be explicitly defined in terms of operations. This error relates to the notion that a theory must be testable. A theory must be testable, but knowing how to test a theory is not equivalent to knowing what it means. Operationism is apparently based on a confusion of these two notions.

The second error, a syntactical one, is the assumption that all terms in a theory must be explicitly defined. In fact, quite the reverse is true. In highly sophisticated theories many basic terms are simply defined implicitly or contextually by the role they play in the postulates of the theory. Such terms, which designate hypothetical constructs (entities or processes which are unobservable), are called theoretical terms. The inclusion of some such terms is a logical requirement for any theory which will combine a high degree of explanatory power and predictiveness with the characteristic or parsimony or elegance. I shall explore this theme in greater detail in the section of this chapter entitled "The Requirement of Theoreticity", below

For the moment we are left with the following question: Why did Bridgeman arrive at this analysis? He was, after all, a practising physicist of considerable repute. The answer, I think, lies quite simply with the circumstance that he was an experimental rather than a theoretical physicist. For many years he was in charge of the Harvard laboratories where he supervised pioneering investigations of the properties of matter under high pressures. (His experimental investigations included the determination of the electrical and thermal properties of various substances at pressures as high as 100,000 atmospheres.) The nature of his work would account for the preoccupation with testing and operational procedures that is reflected in his analysis of science.

#### *The Requirement of Theoreticity*

Let's now take a closer look at this assertion I have made to the effect that successful science is necessarily discourse which features a certain degree of theoreticity. In this context it is useful to consider some work which emerged from the approach to studying the characteristics of science which is sometimes called logical reconstructionism. Essentially, what is meant here is the attempt to uncover, using techniques of formal logic, set theory, and model



theory, the logical or structural features of scientific discourse. Logical reconstructionism got under way in the late 1920's and was originally associated with the Vienna group which formulated the fundamental doctrines of logical positivism. These techniques are still employed and their apparent usefulness has therefore outlived the popularity of the substantive views associated with logical positivism.

There are many themes central to logical positivism, but the one of concern here is the belief that theoretical terms are dispensable, that from a cognitive point of view they are just so much salad dressing. The history of the philosophy of science is characterized from the late 1930's to the late 1950's, by a concerted effort (a research program, if you like) to substantiate this thesis. The climax of this research program came in 1956 when William Craig proved a certain theorem in the first-order logic. The article in which this result was published was entitled "Replacement of Auxiliary Expressions" and it constituted an effective procedure for reaxiomatizing a theory in a restricted vocabulary (Craig, 1956).

The logical positivists were quick to seize on the apparent implications of this work (cf. Hooker, 1968). Craig's theorem entailed the following: For any standard theory,  $T$ , which features an observational vocabulary and a theoretical vocabulary, it is possible to generate a new theory, call it  $T^*$ , which has exactly the same empirical consequences (i.e., has the same theorems in the observation language) as  $T$  but which contains as extralogical terms only the observational terms featured in the original theory,  $T$ .

This seemed to demonstrate that theoretical terms were dispensable. But, in fact, quite from making the point, this result spelled the end of logical positivism. How so? It showed just what one got in the way of a "scientific theory" when one excluded theoretical terms. The first feature which emerges from a study of Craig's result is that while a standard theory generally has a small finite number of axioms, its Craigian counterpart will generally exhibit an infinite number. A Craigian theory is thus hopelessly unwieldy, impossible to either apply or to understand entirely. In the second place, there is no way to acquire a Craigian theory except by first possessing the standard theory from which to generate it by Craig's technique. The gist of all this is that Craig's result had the effect of demonstrating the indispensability of theoretical

terms by showing just what one got when one asked for a science without theoretical terms.

Where does this leave us? I think we must conclude the following. Even if experimentation as we currently understand it in the field of educational research is continued for a substantial period of time, the results will necessarily be limited. Limited but still important, perhaps. The history of the hard sciences reveals a long period during which low level empirical generalizations were apprehended, prior to the emergence of truly theoretical science. But what we need to do, I think, is to begin looking forward to the development and testing of theoretical frameworks in their true sense. We need to purge our field of the aversion to hypothetical constructs.

The conditions for accomplishing this are ripe in at least one respect. One of the obstacles to this development in the past has been the limitation of our mathematical models. The standard regression and analysis of variance techniques are not amenable to the elaboration of complex theories featuring hypothetical constructs with complex patterns of association among themselves and with respect to observed variables, including some terms which will be only implicitly defined. There do exist, however, more sophisticated models in the social sciences -- for example, the structural equation framework (Bentler, 1983; Asher, 1981) -- which can be used in the elaboration of such theories; and, what is more important, they are becoming implementable from a practical standpoint as a result of the development of software packages such as LISREL (Long, 1983). With mathematical models of this sort it is possible to represent theories which feature hypothetical constructs ("latent" variables, in the jargon) in which these variables are regressed on to one another; we are not limited to covariances as we are in factor analytic models. Theories formulated with these models may also feature reciprocal causation (that is, they need not be recursive).

Unfortunately, there is little to be gained from the availability of such sophisticated modelling tools if they are coupled with a poor appreciation of the nature of theory. To illustrate this point consider the following case from the literature of educational research.

Schneider and Trieber (1984) published an article in the *American Educational Research Journal* in which they utilized a latent variable causal modelling approach to investigate the

dynamics of the process of student (math) achievement. The theoretical model they test incorporates three latent variables: Instruction, Aptitude and Achievement. The design follows a multi-wave longitudinal configuration: student math achievement is measured on four occasions over a one-year period. There are a number of problems apparent in the study that relate to the following issues: (i) the theoretical model selected, (ii) the quality of the data in relation to the level of the data required for the statistical machinery employed, (iii) the confounding of exploratory and confirmatory activities, and (iv) the reasoning underlying some of the conclusions which are based on the results. (i), (iii) and (iv) are of particular interest to us here in our concern with the understanding, or lack of understanding, of what constitutes good theory and problems associated with the absence of rigour in the logic of our schemes for testing theoretical conjectures.

In the first place, the study purports to be confirmatory in nature: the authors set out to test a prespecified model of the achievement process against two subgroups of classrooms which apparently feature different instructional histories. The hypotheses are not so clearly stated as one would expect of a confirmatory study. Nor, indeed, are they easily inferred. The following (implied) hypotheses can be identified:

- (1) Different solutions to the measurement and/or structural components of the model will be required to fit the two subgroups.
- (2) The general model will be confirmed.
- (3) Aptitude and Instruction will be uncorrelated.
- (4) The behaviour of a submodel, the dependent variable block, will correspond to a simplex structure.

The subgroups featured in this study were distinguished on the basis of the extreme slopes of their pre- and posttest achievement regression equations. ("HiSlope" versus "LoSlope"). The researchers were unable to fit the data to the *a priori* model in the case of the LoSlope group. For this group, the process variables chosen to represent the latent construct called Instruction did not even successfully define a common factor. A greater degree of success was encountered in the case of the HiSlope group, and the authors assert that the full

model estimation produced for this group represents "an empirical confirmation of the theoretically assumed model" (p. 207). Elsewhere they write: "there is strong evidence that the proposed model described and explained educational achievement in mathematics for the HiSlope sample". These conclusions certainly require some comment, for they are not entirely warranted.

To begin with, this study did not truly achieve statistical confirmation of the implied hypotheses. Hypothesis (4) was confirmed in a separate test directed specifically to the dependent variable block in the case of the HiSlope sample. However, in the jump to the full model specification this assumption had to be abandoned. Furthermore, in the full model the exogenous variables did not contribute to the achievement variation measured on the last two occasions. Thus, while some features of the *a priori* model -- for example, the independence of the exogenous constructs -- were confirmed, others were not. The claim, then, that the *a priori* model received an "empirical confirmation" is subject to important qualifications. Moreover, it is a claim which is based in judgement, rather than one which is founded explicitly on statistical hypothesis testing.

Within the context of the process of elaborating models, the significance of the partial confirmation perceived by the authors of this study needs to be addressed further. The claim is made that, in addition to providing support for the simplistic theoretical model advanced in the study, the results obtained also support the assumption of the intrinsically "local" nature of the achievement models -- an assumption embodied in hypothesis (1). So far as the enterprise of elaborating and testing models is concerned, the results might as reasonably be construed as pointing to the inadequacy of the model presented, and the need to develop a more elaborate model which might explain the behaviour of both samples successfully. Ironically, the authors themselves point to the need to create more sophisticated models. One specific suggestion they advance concerns the possibility of revising the present model to include a multi-wave collection of aptitude data. The failure of the simplex assumption for the full model estimate may point to a reciprocal link between aptitude and achievement not identified in the model tested.

All the same, Schneider and Treiber seem oblivious to the following considerations. The

goal of causal modelling is to produce the most elegant model which is compatible with a certain level of explanatory power. In the present case, even if the proposed model had achieved a complete confirmation in the case of the HISlope sample, it would still only serve to explain the dynamics of the achievement process for four classes out of an original sample of 113. Parsimony is a prime desideratum of latent variable causal modelling. But so, too, is explanatory power and scope and there should be limits on the extent to which we are prepared to sacrifice these to elegance. A model which achieves confirmation for less than four percent of the cases encountered is of limited value and interest. The authors stress, and here is the crux of the problem, that the failure of the model to generalize to both subgroups confirms the need for "local" models of achievement. It may equally be said to point to the need to develop more elaborate models which might effectively encompass a greater variety of cases. It might be possible to elaborate a number of local models with distinctive characteristics which would enable us to describe and predict the behaviour of various different subgroups. But ultimately the goal of causal modelling must be to produce more general models with proportionately greater explanatory power. The point that needs to be emphasized is that the failure of a model to generalize across different samples should not necessarily be interpreted as evidence of the irrevocably local character of a process, particularly when the assumed model is as simplistic as the present one appears to be.

This, of course, is one single study selected from a leading journal; but I believe it is characteristic of certain tendencies in research in our field insofar as it (a) exhibits rash or unwarranted conclusions concerning what precisely has been confirmed, and (b) betrays a lack of appreciation for the real goals and desiderata of theoretical conjecturing.

While I have taken this example from the literature of educational research, it is not hard to find similar examples from the publications perhaps more closely identified with educational technology. A good illustration is Sweet's recent work with the same mathematical tools (path analysis or causal modelling) in the context of distance education research. Sweet (1986) attempts to specify in greater detail an earlier model suggested by Tinto (1982), in order to explain course dropout. A long list of variables are included: locus of control, academic

performance, grade expectation, goal satisfaction, demographic variables, course materials, sex, age, tutor rating and re-enrolment. The end result is a model of "course persistence" with an R-square of .19.

There are several criticisms which can be raised in conjunction with this study: for one thing, it attempts to apply a model of program discontinuation to course dropout (the phenomena are not the same, surely) and, for another, it does not include data collected during the duration of the course, so that it is difficult to see how directional causal relations can be inferred. But, more strikingly, there is the problem of the size of the R-square: with such a small amount of variance explained, in the last analysis this complex model seems to be a case of much ado about nothing. There is a lesson here which I will explore in more detail presently, and it is this: we accept models that lack explanatory power and precision, and our methods of testing are a contributory factor to the vagueness of the theories which are developed.

The increasing use of more sophisticated mathematical apparatus is an encouraging trend in educational research. (Only fifteen years ago, most studies were conducted using univariate statistical tools to test hypotheses that really required multivariate models.) Mathematical models such as the structural equation framework still have their limitations, their associated problems, but they represent a step forward. There will ultimately be a price to pay for this step. We may have to abandon favourite themes such as linearity and recursiveness, although these are dubious in any event. If a successful science of educational processes can be elaborated, the cost may be a requirement of increased mathematical sophistication on the part of the researcher/theoretician. This will increase the gap between the researcher and the practitioner in certain respects, though the practitioner would likely trade reliable generalizations with some meat attached to them for a decrease in the accessibility of the research that might lead to such principles.

We would do well, though, to remind ourselves that the appropriate analytical tools and the skills to use them are only a necessary condition of scientific advancement. What is even more crucial is the identification of key variables of interest and a good initial qualitative understanding of the dynamics of their relations from which to proceed.

*Routing Out False Conjectures: The Limitations  
of Social Science Methodology*

Let us now turn our attention to the testing of theory, and the problem of the lack of rigour invested in the procedures utilized in social scientific research for this purpose. Recall that earlier I raised the point that the methodology of science is self-correcting, implying the importance of procedures that are effective and efficient at routing out false conjectures. A fundamental shortcoming of the methodology of the naturalistic approach as it has evolved in the social/behavioural sciences is the very lack of procedures that are capable in this regard. There are actually several difficulties that can be identified with soft science practices as regards testing of theory. We have just witnessed one: in the Schneider and Treiber study one part of a model fit a very small part of the sample, but this was taken to be essentially a confirmation of the model (once several *ad hoc*, mitigating, considerations were brought to bear). I shall return to this particular type of methodological sleight of hand, but it would be best to begin with a more fundamental, deep-seated problem which is at the core of a network of interconnected difficulties that will be laid out in this section. This problem is referred to in the literature as Meehl's paradox after Paul Meehl, a clinical psychologist, whose seminal article entitled "Theory Testing in Psychology and Physics: A Methodological Paradox," (1967) drew attention to the obstacles placed in the path of falsifying theories and hypotheses in soft science research.

Meehl was rightly struck by the differences in the approaches to theory testing exhibited by psychology and related "soft" sciences as opposed to physics. I have already touched upon how theory testing proceeds in the hard sciences via the hypothetico-deductive scheme. Recall that a theory, T, together with certain auxiliary conjectures (e.g., theories concerning how the instruments that are deployed operate) and information (definitions, rules of correspondence, statements of initial conditions), yields as a deductive consequence a prediction concerning a specific event. Observations and measurements are made under controlled (experimental) conditions and the results of these are compared with the prediction. If they fail to square with the prediction, then the whole apparatus mentioned above is brought into question by the logic of the *modus tollens* inference rule. The theory, or some part of the theory, or possibly some

part of the auxiliary component which figures in the derivation of the prediction, must be assumed to be false. Note three things: first, the prediction is a specific one which names an exact predicted value (or possibly a particular form function) of the experimental variable of interest. Secondly, the predicted value is derived logically from the substantive theory concerned. Finally, there is an asymmetry between the force of a corroboration and that of a failure of the prediction. Corroboration may increase our faith in a theory, but can never do so beyond the point of a conditional acceptance, for any number of successful predictions is compatible with a theory being false in some respect. On the other hand, a recalcitrant fact is decisive, as previously noted.

None of these three characteristics is present in soft science theory testing. The approach used in the soft sciences is not based on point predictions or exact expectations of the values of experimental variables. Instead, a familiar procedure called null-hypothesis testing is employed. Here, essentially, the means obtained from that portion of a sample assigned to a treatment condition are compared to those assigned to some other treatment conditions and/or a to control group. In the simplest case, one treatment condition contrasted with a control group, the null hypothesis, expressed in the form  $H_0: |u_a - u_b| = 0$ , states that there is no difference between the two groups other than what one would expect based on random variation. If a difference is found between the two groups that is greater than what is attributable to chance then the null hypothesis can be rejected, and *indirectly* the alternative hypothesis (asserting that there should be a difference and that this is attributable to the particular manipulation invoked for the independent variables) is confirmed.

Such "confirmation" is rather more oblique than the sort we encounter in physics, and its force is correspondingly weaker. In fact, as Meehl argued, null hypothesis testing hardly provides any genuine test of the mettle of a conjecture at all. The problem lies with the counterproductive effects of any increase in the proficiency of the instrumentation employed in research. In the case of physics, any increase in the sophistication of the instruments, which leads to more exact determination of the value of experimental variables, results in a more stringent test for the theory concerned. The point predictions yielded by the theory will have to



become increasingly exact in order to fall acceptably close to the actual measurements recorded. But, oddly, any increase in the accuracy of instrumentation in the context of the soft sciences only increases the likelihood that the null hypothesis will fail (that a difference, for whatever reason, will be found) and thus that the experimental hypothesis, H, will be indirectly confirmed. Given a large enough sample, the non-directional null hypothesis ( $H_0: |u_a - u_b| = 0$ ) will be "quasi-always" false, while a directional hypothesis ( $H_0: u_a - u_b \leq 0$ ) or ( $H_0: u_a - u_b \geq 0$ ) will be false at least one-half the time, and more realistically will be rejected to some undetermined degree greater than one-half of all occasions. This perverse effect of improvements in instrumentation is called, somewhat unsuitably, Meehl's "paradox". In effect, it entails that any hypothesis, no matter how doubtful, will be confirmed at least half the time, given a large enough sample and sufficiently precise instrumentation.

Meehl's attack on soft science methodology has not been parried with any success in the intervening two decades, though there have been responses. One of the most detailed has been put forward by Serlin and Lapsley (1985), who have argued that his critique is based on a simplistic, naive form of Popperian falsificationism which undermines his case for the greater power of theoretical tests in physics. They point out, as I have above, that it is not the theory in isolation which is tested, but rather a whole complex web of assumptions and theoretical components, so that a false prediction is no more fatal to the theory from which it was generated than is a "no significant difference" result (forcing us to accept the null hypothesis) in soft sciences. This defence is not very convincing, however. In the first place, in physical science it may be possible, in some instances, to isolate which portion of the web is to blame in quite decisive fashion. If all theory and auxiliary assumptions are specified with sufficient precision, it may be possible to isolate the portion to blame by model theoretic means (see comments in Chapter 5 regarding independence proofs), providing the different elements concerned are sufficiently independent of one another.

Now it might be argued that in the physical sciences this situation of logical independence is not often the case. While the postulates of a given theory may be mostly independent, the auxiliary theory and assumptions required to render the theory testable are

often not entirely independent of the theory under scrutiny itself.

A striking example of this is to be found in Gauss' attempt to determine the geometry of physical space (Reichenbach, 1951). His observations involved the triangulation of three distant mountain peaks. The idea was that any departure from the Euclidean angular sum of 180 degrees would establish empirically that physical space is non-Euclidean. Unfortunately, the use of optical instruments to carry out this investigation necessitated certain assumptions regarding the behaviour of light rays. A deviation in the measurements obtained from the Euclidean case could be interpreted alternatively as follows: space conforms to Euclidean requirements, but the path of a light ray does not conform to a Euclidean straight line. At this point one might object that, in principle at least, one could verify the assumption concerning the behaviour of light rays by independent means, using measuring rods to check against distances arrived at with optical instruments. But this procedure itself requires a conventional stipulation regarding the behaviour (the rigidity) of transported rods. Thus, we find that conjectures concerning the characteristics of physical space ultimately are sewn up with assumptions of optical theory that dictate our understanding of the instrumentation used to test these conjectures, and these assumptions are not independently testable.

But observations of this kind simply play into Meehl's hands (Dar, 1987). If, typically, the auxiliary assumptions in physics have links to, or are integrated to some extent with, the theory tested then a recalcitrant fact does carry a great deal of weight; it is not so easy to deflect blame away from the theory itself onto the auxiliary hypotheses. So even if, as Serlin and Lapsley argue, a test of a theory in physics is really a test of a complex of theories and assumptions, the extent to which these are interlocked can render a test a serious hurdle. Unfortunately, in the soft sciences the auxiliary hypotheses are rarely closely identifiable with the theory tested (assuming there even is a theory behind the experimental hypothesis). In particular, instrumentation in soft sciences is much more likely to be independent of the theory being tested, and to be open to question, than is usually the case in physics. It is almost always possible, logically speaking, to blame a failure to reject the null hypothesis on the instrumentation employed without implicating the theory. From a practical standpoint as well this avenue is often

available since, as Dar remarks, it is difficult to think of any construct in social sciences for which there exists universal agreement regarding what constitutes the best measurement tool.

This lack of interconnectedness or interdependence among elements of substantive theory, auxiliary assumptions and instrumentation, in conjunction with null hypothesis testing enables several of the pernicious "social forces and intellectual traditions" Meehl identifies in soft science research: One of these is a tendency to dilute reports of research findings with liberal dosages of *ad hoc* explanations intended to mitigate unanticipated results – there is an element of this in the Schneider and Treiber study outlined earlier. This leads to facile circumventions of *modus tollens* refutation, and undermines the goal of achieving powerful research programs. Meehl is quite forceful on this last aspect:

It is not unusual that . . . this *ad hoc* challenging of auxiliary hypotheses is repeated in the course for a series of related experiments . . . In this fashion a zealous and clever investigator can slowly wend his way through a tenuous nomological network, performing a long series of related experiments which appear to the uncritical reader as a fine example of "an integrated research program," *without even once refuting or corroborating so much as a single strand of the network.* (1967, p. 114)

Another discrepancy between the soft sciences and the natural sciences relates to the derivation of the hypothesis tested. In physics, the experimental hypothesis is a deductive consequence of the substantive theory. So the full force of the *modus tollens* rule is in effect. On the other hand, the null hypothesis tested in behavioural research is not derived from any substantive theory at all: it is merely a methodological device.

This has a bearing on Serlin's and Lapsley's proposal for modifying null hypothesis testing to approximate the strength of point prediction schemes. Rather than state the null hypothesis as ( $H_0: \mu_a - \mu_b = 0$ ), they recommend the following formulation: ( $H | \mu_a - \mu_b | \leq \#$ ), where  $\#$  is the difference that we have decided beforehand we will accept as one which is truly due to experimental treatments.  $\#$  is presumably based primarily on an estimate of measurement error given state-of-the-art experimental technique. The idea is similar to what we encounter in physics, where a so-called "good-enough" belt is established around the

expected value of a variable, specifying exactly what differences between the theoretical prediction and the experimental result are to be considered tolerable.

However, Orey, Garrison and Burton (1989) argue that the good-enough belt proposal does not succeed in bridging the gap between the methodological practices of soft sciences and physics. The problem they identify is that in physics the value for  $\#$  is constructed *ex post facto* rather than *a priori*, as would necessarily be the case in Serlin's and Lapsley's emendation. In physics  $\#$  is essentially an estimate of measurement error that reflects the limitations of the instrumentation employed.

To be more precise, there are actually two basic kinds of error to be considered. The first category, random error, result from fluctuating conditions and small disturbances. These are said to be "determinate", meaning that they can be assessed by some logical procedure (usually the statistical theory of error). The second, called systematic error, is more problematic: systematic error may be indeterminate, i.e., there may be no effective procedure for assessing it. Systematic error includes faulty calibration of instrumentation and errors resulting from flawed experimental technique. The only estimate that can be given for the indeterminate component of experimental error is one based on state-of-the-art experimental technique and existing, established theory. The results yielded by instruments are simply compared to the predictions of well-established theories. Thus, the refinement of our theories and instrumentation proceeds in an iterative fashion, with an established set of instruments and theory serving as the basis for the development and validation of new measures. What is obviously particularly problematic, though, given this recursive form of development, is the initial move in theorizing: the jump from qualitative to quantitative concepts. The initial construction of a fundamental scale must occur without any foundation of established science to serve as a stepping stone. Orey et al. seem to feel that this alone rules out the Lapsley and Serlin proposal (pp. 19-20).

There is something to this objection. After all, we are, in the soft sciences, at this very stage of progressing from qualitative to quantitative thinking. We are not sure what objects are fundamental, we are not certain what to measure, and our instruments are highly problematic from the point of view of construct validity (except where a kind of premature reification has

occurred as in the case, for example, of IQ – but this is evidently not desirable either). However, any science must go through this initial stage, and history teaches us that lack of precise estimates of errors of measurement and errors of observation should not be used as an excuse for foregoing point predictions. Modern astronomy is one area within science in which the systematic analysis of error is highly developed. Yet its precursor, the Ptolemaic theory, had no explicit theory of error – even though, despite the impressive precision of this theory and the goodness of fit it achieved with the data, the relative crudity of observation methods of the time guaranteed discrepancies between the theory and observation. Indeed, even the most impressive accomplishments of classical physics, Newton's and Maxwell's theories, do not contain any systematic treatment of error.

Admittedly, these theories are all of the strictly deterministic variety, and this is not the form we generally suppose soft science theories must assume. But it is not obvious that there is anything to be made of this contrast in the present context. Quite the contrary: in a probabilistic theory it may be possible to simply allow the theory itself to take up the slack between theory and observation (if the discrepancies are relatively small). The result may certainly be more elegant than what can be accomplished with a deterministic theory where the treatment of error is not well integrated. In fact, two external mechanisms may be required in the deterministic case. A theory of measurement error must first be advanced before the predictions of the theory, compensated by the estimates of error, can be tested. And, further, even at this point it may still be necessary to decide whether small discrepancies that are left unaccounted for can be ignored.

Orey et. al.'s objections to point null hypotheses based on difficulties associated with estimating error and the necessarily largely *a priori* character of these estimates thus do not go through. However, they do point out, quite correctly, that the point null test is still a test of an hypothesis that is forwarded on methodological grounds rather than deduced from substantive theory, and that consequently the proposal simply resurrects Meehl's paradox. As instrumentation becomes progressively refined, this must be reflected in a lowering of the value assigned to  $\alpha$ , until  $\alpha$  tends to zero. However, as  $\alpha$  approaches zero, the point-null hypothesis

simply becomes the conventional null hypothesis and is, once again, quasi-always false (p. 20)

The conclusion to be drawn from the foregoing discussion, I think, is that null-hypothesis testing, as a vehicle for developing and testing theory, has serious limitations and should probably be abandoned. The case for this position can be made even stronger by considering, to start with, the remaining, associated problems identified by Meehl. The first is that there is a tendency to conflate the statistical hypothesis tested, which is strictly a methodological device, with the substantive theory from which the alternative hypothesis is presumably derived (1967, p. 107). An implicit belief is thus often formed, one which has no basis but which is insidious, that the level of statistical significance is somehow a direct measure of the probability that the research hypothesis is true. The second is a proclivity for "counting noses": disregarding the logical force of *modus tollens*, social scientists tend to treat disconfirming instances on a par with confirming instances so that if a theory has more confirming than disconfirming ones then the evidence for it is generally taken to be relatively good. The error of this methodological game is further compounded by the circumstance that, even if the practice had some merit in principle, one cannot even really establish the proportion of disconfirming cases since nonsignificant results are generally either not reported in published form, or else are discounted and explained away on *ad hoc* grounds.

Cohen and Hyman (1979) have defended null hypothesis testing on nonstatistical grounds. Their argument is that experimenter bias constitutes the greatest threat to internal validity, and that this problem is in large part an attitudinal one which cannot be dealt with through explicit technical or statistical aspects of research design: "somehow, attitudinal realities appear to become tangible but 'invisible' realities causing extraneous variance. In other words, a large part of research design is an attitude of objectivity" (p. 16). One way to minimize this problem, they maintain, is to follow a tradition that demands that we prove nulls rather than our research hypotheses. Thus, their defence of the null tradition rests on the claim that it promotes greater objectivity on the part of researchers. Clearly, though, there are alternatives. One is to train researchers to be more objective. Certainly the physical sciences do not seem to suffer from the circumstance that direct tests are made of research hypotheses. Indeed, the most

blatant cases of experimenter bias (outright cooking of the books) are to be found in the social sciences and medicine. There simply does not seem to be any empirical evidence that directly testing a null hypothesis, rather than a research hypothesis, provides any better guarantee of objectivity. If there were, then innovations such as the development of double-blind procedures would hardly have been considered.

There are still other problems which need to be mentioned. One is that the reliance on tests of statistical significance removes any pressure to develop a precise theory. The way researchers are trained in the social sciences reinforces this: we are taught to be satisfied when a particular statistic is significantly different from zero at the .05 level. As Dar emphasizes, once this convention is accepted not only as a conventional criterion for acceptability for publication, but also as a criterion for successful prediction, then there is no motivation to seek more powerful, integrated and predictive theoretical explanations. This is reinforced by the fact that the sophisticated statistical machinery that has evolved around null hypothesis testing can lend an aura of rigour and scientific credibility to research programs that are otherwise based on weak theories and poorly conceptualized questions.

To summarize: evidently there are many problems associated with null hypothesis testing, but they are all clearly related to two central considerations: First, null hypothesis testing is not an effective procedure for falsifying inadequate, incorrect conjectures; it does not provide the basis for a self-correcting methodology of inquiry that might lead to the cumulative development of knowledge. That researchers themselves in fact do not take the conventions involved in tests of significance seriously is also apparent. Few skilled investigators would repudiate a favoured hypothesis solely on the grounds that it only achieved significance at, say,  $p = .06$ . Second, it obviates the requirement for precise, highly consolidated theoretical frameworks.

### *The Formative Evaluation Approach to Theory Testing*

While these problems are endemic to the social sciences in general, and hence can be generalized to practices within the field of educational technology, I shall conclude this chapter with some comments concerning a recent suggestion regarding the testing of theory which has

arisen specifically in the literature of educational technology.

Reigeluth (1989), in a position paper concerning new directions for the field, has suggested a strategy for research that he labels the formative evaluation approach. The basic idea in this proposal is to complete the design and development of an instructional product solely on the basis of an instructional theory or model that you wish to research. "Try not to use any other prescriptions, not even your own intuition in designing the product" (p. 71). The second element concerns the notion that one must "use theory to design the product . . . while noting any instructional features you use which aren't prescribed by the theory" (p. 72). In the final step, the product is evaluated and the results are reflected back upon the underlying theory or model of design.

There are serious problems associated with these ideas as a basis for empirical research and theoretical development that go beyond anything I have previously raised. There are, so far as I can see, three quite transparent difficulties.

(1) The first is that no "prescriptive" theory or model of design in the instructional arena is sufficiently detailed to permit the completion of an artifact without resorting to intuition or other generalizations external to the focal model or theory. This is implicitly recognized in step two of Reigeluth's proposed methodology, where he emphasizes the need to detail those aspects of the design which are not determined by the model or theory. But then the very idea that the success or failure of the product can be taken as any meaningful confirmation or refutation of the underlying prescriptive theory is incoherent. The blame for any failure in terms of the evaluation results can be shifted onto the features incorporated which aren't included in the theory, or onto the intuition or creativity (or lack thereof) of the researcher/producer. Similarly, success may reflect the power of the generalizations comprising the theory, or it may conceivably owe primarily to the intuition or external principles brought to bear.

(2) There is also the question whether prescriptive theories are exclusive or unique enough for products designed and developed accordingly to furnish "crucial" tests; that is, tests that can decide between what are ostensibly competing theories. This an issue made even more moot by Reigeluth's insistence that prescriptive theories, as opposed to the "descriptive"



theories, as he refers to them, upon which they are allegedly built, must be eclectic.

(3) This also raises the question what exactly is being "tested" in formative evaluation research. Is it the prescriptive theory or the underlying empirical theory. If it is the underlying theory, then why confound the matter by addressing it indirectly, through some associated prescriptive theory? The answer is perhaps apparent in the latter pages of Reigeluth's article, where he draws attention to the fact that in the United States, at any rate, more funds are now available to formative evaluation research (in the form of development funds) than can be tapped for more purely empirical, theoretical research.

I do not wish to suggest that Reigeluth's entire proposal is based on matters of expediency associated with funding. All the less so because Clark (1989), a figure who has been much preoccupied with the quality of research and with methodological questions in our field, has also supported this direction. Clark's enthusiasm, though, is based at least in part on his contention that there is not enough research that compares one model or theory of design to another. Thus, Clark's interest in the idea seems to be predicated on its role in assessing applied or prescriptive theory, while it is not so obvious that Reigeluth shares this view. Reigeluth does not make his views concerning the relationship between what he calls descriptive theory and prescriptive theory explicit, but one can infer that he conceives of the relationship as a "loose" one. His contention that prescriptive theories need to be eclectic, while descriptive theories should not be, suggests this. If prescriptive models and theories are to follow closely from some underlying descriptive theoretical foundation, and if a high degree of consolidation and integration is expected of this foundation, then most likely one would expect and demand these characteristic of the prescriptive discourse as well.

So there is some reason to believe that Reigeluth, too, believes formative evaluation research is primarily directed at the corroboration or falsification of prescriptive theories. But then we find ourselves on the horns of a dilemma of sorts if we try to defend his proposals. Either formative evaluation research is aimed primarily at assaying prescriptive theory, in which case it is by hypothesis no replacement for basic research aimed at developing descriptive theories, or else it is intended as a replacement for such basic research, in which case it is an

approach that is so indirect as to be indefensible. What we need, it must be reiterated, is stronger tests of our ideas, tests that will eliminate poor conceptualizations and flawed conjectures, and force us to develop more precise, more powerful explanatory hypotheses. Reigeluth's suggestions do nothing to improve our lot in this regard and, given the foregoing discussion, might even be regarded as a regressive step.

One can question, also, his assertions regarding the desirability of eclecticism in prescriptive theory. One explanation of why this seems to be a desirable quality may simply be that the descriptive theories extant which might inform the content and structure of tightly integrated prescriptive theories are too incomplete and too weak, too immature, to perform this job. (Another is that many of our prescriptive theories or models of design are not really based on descriptive theory in any significant degree (Andrews & Goodson, 1980)). This is not necessarily a criticism, since it is not self-evident that well-developed basic empirical theories -- of learning, or of development, for example -- will necessarily translate into highly effective, precisely articulated prescriptive instructional theories or "technologies of instruction".)

#### *Conclusion*

The likelihood that we will meet the challenge, forego null hypothesis testing and abandon any strong commitment to operationalism is perhaps slim. The problem, in large part, is that operationalism and null-hypothesis testing are easy to teach, in a highly proceduralized fashion. Research methods texts do not really try to address the issue of what constitutes good, solid, explanatory theory. They tend to reflect naive operationalist tendencies and to limit their scope largely to statistical issues.

A brief survey of basic research methods texts in education will bear these observations out. Slavin (1984) refers briefly to the notion of "theory" on one page (p. 7), and the term does not appear at all in the index. Borg and Gall (1983), a widely used text, introduces the notion of a "theoretical construct", and explains that such concepts can be defined constitutively (by specifying their relationships to other concepts) or operationally (by reference to measurement procedures). But, interestingly, the example of a constitutive concept which is provided -- the Piagetian notion of conservation -- is one which is readily definable operationally. A more recent

edition (Borg & Gall, 1989) is notable for its inclusion of some discussion of principles pertaining to postpositivistic philosophy of science, including criticisms of the fact-value distinction, incommensurability and underdetermination. Unfortunately, these principles are presented in an uncritical manner; the counter arguments are not raised. Linn (1996) contains no mention of theory or operational definitions and their role. McMillan and Schumacher (1989) discuss operational definitions (pp. 84-86) and mention constitutive definitions of constructs, also. However, there seems to be, implicit in their discussion, the presumption that constitutive definitions must be operationally defined for research. Tuckman (1978) includes a discussion of intervening variables (pp. 67-70) and a chapter entitled "Constructing Operational Definitions of Variables" (pp. 77-91) that presents a typology of operational definitions and the reasons for using them. Bridgman (1927) is a prominent reference. There is no detailed discussion of the meaning of "theory". Johnson (1977) includes a brief discussion on the topic of theory (pp. 36-38), but the examples of "theoretical" propositions which are provided are hardly exemplary: for example, "it is more efficient to learn material as a whole than to break it up into parts," and "additional replications of a task become less efficient for learning and retention." Dyer (1979) contains a chapter on educational theory which investigates the different types of theory (pp. 348-349), but in a book of several hundred pages this chapter is a mere two pages in length. Kerlinger (1973) includes a discussion of operational definitions and constitutive definitions of constructs. He includes a useful distinction between two types of operational definition: measurement versus experimental operational definitions, where the latter concern how an experimenter manipulates a variable. Of the basic research methods texts I have seen, Kerlinger alone does not appear to hold that all constructs can be operationally defined and should be if they are to play any role in the development of theory. However, Kerlinger's text is also the only one mentioned here that is not specifically oriented to educational research.

Again, these observations reflect how operationalism and the logic of null-hypothesis testing are easy to encapsulate and disseminate. And, after all, there are no "paradigm" cases of social science theorizing to present as exemplars in research texts, so that presenting a more sophisticated account of the nature of theory is not a simple proposition. But this is not the

entire appeal of these items. They also provide a very nice, neat set of conventional criteria for deciding on the acceptability of reports for publication or dissertations.

So it is easy to be sympathetic with the situation as we find it, even if we acknowledge the dangers. It is also true that the null approach would be slightly more defensible if replication were common, but it is not (*cf.* Shaver & Norton, 1980). The reward structure of the field simply does not encourage replicative studies. In fact, it is arguable that replication of new results carries greater weight in terms of the acquisition of status for researchers in the field of physics than it does in the field of education. This reward structure is firmly entrenched; one can speculate that altering it would perhaps be more difficult than effecting a move away from null-hypothesis testing altogether.

Null-hypothesis testing would also be more defensible if true randomization were common in research designs. However, as mentioned before, this is not compatible with the logistical and bureaucratic constraints inevitably imposed on studies conducted in natural settings. Random assignment of intact groups to treatment conditions merely increases the likelihood that there will be some prior differences, unrelated to the treatment, which will lead to a rejection of the null hypothesis.

But what is the alternative? If we do not make some move towards adopting a methodology that will support "strong inference", serve to falsify incorrect hypotheses, and encourage the development of truly explanatory frameworks, then the "scientific" approach in educational technology research will deservedly pass into oblivion. The naturalistic approach will be bypassed, without ever having been given a true test.

## CHAPTER 4

### Misconceptions Concerning Theory and Theory Development: Some Illustrations from the Literature

In the previous chapter I argued that while no decisive case has ever been made against the possibility of a naturalistic approach to social science, in principle, certain contingent aspects of the current model pretty well ensure that no significant progress will occur until adjustments are made. Two problems were discussed in some detail: the weakness of null hypothesis testing as a procedure for eliminating false conjectures, and a faulty conception of theory, widely upheld, based quite strongly on the philosophy of operationalism.

In this present chapter I want to pursue the issue of our understanding of theory. It is axiomatic that sophisticated naturalistic theories will not evolve until there is widespread appreciation of their nature and characteristics. What I want to illustrate in the following sections is that our unsophisticated conception of theory is evident not only from the contents of the all too brief sections in research methods textbooks that are devoted to the topic, but is also inevitably reflected in the products of our research: both in the items which are presented that are alleged to constitute theory, and in discussion within the research literature of how theory development in different areas can be informed and guided.

I have chosen several cases which I take to be representative and which, at the same time, are closely tied to the literature and concerns of educational technology. Included are the research programs on visual learning, instructional theory, learning theory, instructional design models and theory, theories of mediated instruction, and theory related to distance education.

The criticisms and observations which will be advanced concern the following points:

1. The term "theory" is usually merely an honorific one, not warranted by the stature of the discourse to which it is ascribed. It is applied to a variety of discursive objects which fail to meet the usual general criteria for theory-hood: predictive and explanatory power, scope, precision and falsifiability. Metaphors often parade under the guise of theory. So, too, do taxonomies. Both may play a role in theory development but they are not, in themselves, deserving of the label. In other cases, low grade empirical generalizations (statements which are couched in strictly observational terms and which consequently are not truly explanatory) and

virtual truisms are cast in the role of theory.

2. There is considerable confusion over the relationship between theories and models (Part of the problem here is the lack of any standard definition of the term "model".)

3. There is also confusion evident over the relationship among different kinds of theories. In particular there is, in the field of educational technology, a tendency to blur the important distinctions among theories of learning, theories of instruction, and conceptual models of instructional design.

4. There is confusion over the process by which theories are developed -- which simply reflects, again, the weakness of the underlying conception of the animal.

#### *Holmberg's Theoretical Framework*

##### *for Mediated Instruction*

Holmberg's (1985) work concerning the search for a theory of mediated teaching provides several clear examples of the misconceptions and confluences mentioned above. These are all the more striking given that Holmberg is at pains to clarify the conception of theory operative in his essay. He distinguishes two initial broad senses. The first owes to Gagne (1963, p. 102) and denotes "any systematic ordering of ideas about the phenomena of a field of inquiry." This is a weak sense of the term which is hardly coextensive with the idea of theory in empirical science. It is also, one may remark, a definition which would allow a taxonomy to qualify as a theory. According to the second sense, theory means "a set of hypotheses logically related to one another in explaining and predicting occurrences." Such theories are said by Holmberg to comprise statements of the form "If A then B" or "the more A the more/less B" (p. 16). It is the latter conception which Holmberg adopts.

From this basic starting point Holmberg proceeds to refine his conception of theory and to make a case concerning what precisely one can reasonably expect a theory of mediated teaching to deliver. From there he continues to present and defend an ostensibly theoretical framework for mediated teaching.

#### *Deductive versus Inductive Theorizing*

Two further distinctions concerning theory which are subsequently introduced are. (a)

deductive versus inductive theorizing, and (b) predictive versus descriptive theory. So far as (a) is concerned, Holmberg suggests that the inductive approach involves the collection and evaluation of data without any background theory guiding the process. While this approach was, he says, acceptable during the ascendancy of behaviourism and logical positivism, it has now given way to a general requirement that "a theory must be developed before any empirical investigations are made in order to guide the investigation and make deductions possible -- a deductive approach" (p. 17).

Holmberg is right to point out that there is at least greater lip service paid today to the notion that experimental hypotheses should be derived from some theoretical framework, but in any other respect the distinction he is proposing between deductive and inductive theorizing is entirely a spurious one. To begin with, the notion that one can simply collect data and then evaluate it, searching for patterns and related hypotheses, is a dubious one. There must be at least some implicit conjecture that guides the selection of data and that influences its interpretation. Therefore, so-called inductive theorizing cannot be distinguished from "deductive" theorizing by the idea that it occurs *in vacuo*. In addition, the title "deductive" theorizing is inapt. All empirical theory is necessarily inductive. At bare minimum it will be inductive because the generalizations which comprise theory go beyond any finite amount of data which might actually be accumulated. A simple low level empirical hypothesis, let alone a complex, full-blown theory, will entail an infinite number of predictions. Moreover, a sophisticated theory will also include theoretical constructs, which designate objects and processes that are unobservable in principle and therefore go beyond what is given directly in experience.

The distinction between inductive and deductive reasoning, so far as scientific method is concerned, is a significant one. But the distinction to be drawn is not one between inductive and deductive theorizing. Rather, the distinction is tied up with the contrast between the logic of justification and the logic of discovery. The creation of a theory is always an inductive act. However, recall that the testing of theory usually proceeds by the so-called hypothetico-deductive approach: a prediction is generated (using deductive reasoning) from the theory combined with certain auxiliary assumptions concerning, e.g., the functioning of the

instrumentation employed and statements of initial conditions, and this prediction is compared with actual observations taken under controlled conditions. Thus, it is the corroboration or testing of theories which is essentially a deductive process, not the creation of theory.

#### *Descriptive versus Prescriptive Theory*

Holmberg also distinguishes between descriptive and predictive theory. He first lists four basic characteristics of theories of teaching: logical consistency; specification of functional relationships between teaching and learning outcomes; derivability of specific hypotheses and predictions, and; falsifiability. He then enumerates four basic requirements of a theory of instruction, as per Bruner (1971): specifying experiences that will motivate individuals to learn, or make them receptive to instruction; defining "optimal structure" in relation to the learner's characteristics, to simplify information and render it more manipulable; specifying optimal sequencing of instruction, and; specifying the nature and pacing of rewards and punishments (Holmberg, p. 17). After presenting these requirements he then says that a theory of teaching is "evidently predictive (technological)" and in this sense is to be contrasted with a theory of learning "which is descriptive in its attempts to explain how learning occurs." He emphasizes that a predictive theory stresses practical considerations, techniques and means over explanation (p. 18).

This distinction between predictive and descriptive theory is not a particularly happy one. The difficulty is that the distinction is predicated on the role of explanation. Descriptive theory, which apparently designates the products of pure empirical science, is said to emphasize explanation, while predictive theory, which is alleged to be more technologically oriented, obviously focuses on prediction. Unfortunately, the notions of explanation and prediction are closely allied. They are, in fact, virtually two sides of the same coin. To say that we have explained some phenomenon is essentially to say that we can successfully derive a prediction of the event in question from some theoretical framework.

The two notions are not entirely symmetric, however. We would generally refuse to say that a theoretical system explains some phenomenon unless it could be used to generate the relevant predictions. However, we will not always accept that a subsumptive law statement



explains the phenomena it can predict. A simple, isolated law statement such as " $f = ma$ " can be used to make predictions, but it is only within the broader theoretical framework of classical physics, which includes this relation, that the phenomena are truly explained. Similarly, low level empirical generalizations (those which have not been subsumed within an integrative theoretical scheme) can be used to predict, but do not, in and of themselves, explain what they predict.

So admittedly prediction is not to be equated with explanation. However, prediction is at least a necessary condition for explanation in scientific discourse. Hence, it follows that one cannot distinguish descriptive theories from predictive ones on the basis of their relative emphases on prediction and explanation: a so-called descriptive theory cannot truly be said to be explanatory unless it is highly predictive.

### *Metaphor as Theory*

These untenable and misleading distinctions – predictive versus descriptive theory and inductive versus deductive theorizing – are enough to alert us that Holmberg's conception of theory is a muddled one. Matters become more serious, however, as he progresses towards the elaboration of his own "theory" of mediated instruction. He begins by mentioning four basic "theories" of teaching identified by Fox (1983). These are: (1) the transfer theory, "which treats knowledge as a commodity to be transferred from one vessel to another"; (2) the shaping theory, which views teaching as the process of moulding students to fit a predetermined pattern, (3) the travelling theory "which treats a subject as a terrain to be explored with hills to be climbed for better viewpoints with the teacher as the travelling companion or expert guide", and (4) the growing theory "which focuses more attention on the intellectual and emotional development of the learner" (Holmberg, p. 19). Holmberg then goes on to say that "Whichever theory a teacher uses to help him/her think about the process, it will affect the strategies he/she uses and it will even colour his/her attitudes to students and to any training programme which he/she undertakes" (p. 19).

These roles may indeed reflect the goals of a theory of instruction, but they are not fulfilled by the "theories" Holmberg advances. These so-called theories are simply metaphors which describe different ways of looking at teaching and learning. They do not "explain or

predict instructional phenomena", nor do they prescribe methods of instruction, as one would rightly demand of a theory of instruction (Reigeluth, 1983).

There is no question that metaphor can serve a heuristic purpose in the course of theory development. Certainly, the role of analogical reasoning in specific acts of creative scientific thinking (in terms of theory formulation and concept formation) is well-documented and widely discussed in the literatures of the philosophy and history of science (cf. Hesse, 1963; 1966) and cognitive psychology (Bruner, Goodnow & Austin, 1956). However, metaphors are not theories. They lack specificity and they do not provide precise predictions which can be held up against experience to determine the exact extent of their validity.

#### *Common Sense Knowledge as Theory*

After appraising the general groundwork concerning the concept of theory put forward by Holmberg, it is not surprising to discover that his own proffered "theory" is hardly such, at all. He begins by identifying seven basic background assumptions. These fall into three general categories. The first assumption is that teacher-learner interaction can be supplanted by pre-produced courseware which causes learners to contemplate different viewpoints, consider different approaches or solutions and "generally interact with the course" (p. 20). The last is that the effectiveness of teaching is measurable by learning outcomes *vis-a-vis* what is taught. The intervening five assumptions all essentially concern the relations among learning pleasure, motivation and learning. For example, his third assumption asserts that learning pleasure is a factor in student motivation, while the fifth assumption holds that motivation facilitates learning.

These are clearly fairly low-grade premises, particularly the five which comprise the last category. Let us now consider the "theory" which is formulated on the basis of these assumptions:

Mediated teaching will support student motivation, promote learning pleasure and effectiveness, if it is provided in a way: felt to make the study relevant to the individual learner and his needs; creating feelings of rapport between the learner and the course developer and tutor, if any; facilitating access to course content; engaging the learner in activities, discussion and decisions; generally

catering for helpful real and simulated communication to and from the learner.

(p. 20)

Holmberg maintains that his theory has explanatory value in so far as it relates teaching effectiveness to the effects of certain affective factors (feelings of belonging and cooperation) and certain types of (mediated) communicative interaction (p. 20). He also claims that it satisfies the additional general criteria for theories: It is logically consistent, posits functional relations between (mediated) teaching and expected outcomes of learning, and is falsifiable (p. 21). The difficulty is that formulations of common sense knowledge can share all of these characteristics with scientific discourse. The difference between the two lies firstly in the degree to which these characteristics obtain, and secondly in the integrative function of theoretical explanation. One of the chief characteristics of scientific theories is their ability to provide very concise, highly systemic explanations of a variety of phenomena which, from a common sense viewpoint, may appear diverse and unrelated. It seems evident that Holmberg's explanation of the success or failure of mediated instruction (in terms of the factors alluded to above) is simply a common sense one.

#### *Tautologous Propositions as Theory*

In fact, there is a temptation to go a step further and venture that his explanatory framework is virtually a set of truisms. While the components of his theory constitute empirical propositions, and as such are in principle falsifiable, they actually risk very little. There is an inverse relationship between the risk a logically coherent theory takes of being falsified, and its empirical strength: the more a theory says about the world, the greater the chance that it will be mistaken. Holmberg's theory expresses very little and hence runs no great risk. Holmberg denies explicitly that his theory is platitudinous, but his protestations are ultimately not very convincing.

Consider, for example, one of the consequences identified of this theory, namely, a recommendation of "a style of presentation that is easily accessible; a high degree of readability of printed course materials" (p. 20), and the associated research hypothesis: "The more easily accessible the preproduced course (the more readable the texts), the better the outcome of

learning" (p. 22). One linguistic trick to assess the content of an empirical statement is to restate it in negative terms: to stipulate what is being denied, in other words. In this last instance, what is being denied is that making texts more readable will not facilitate learning. But who would ever consider advancing the thesis that making texts more accessible to the learner will have no effect on learning? The only thing interesting about the entire question is how, precisely, to operationalize the concept of "readability".

Holmberg is not alone, of course, in trivializing the notion of an explanatory theoretical framework. In an even more striking instance Beckwith (1989) argues that there is an "implicit unifying theory of educational technology." I quote only some of the principles of this theory "communication of clearly envisioned desired performance facilitates performance attainment and evaluation" and "environments can be structured, in a systematic fashion, to effect successful learning and motivation" (p. 128). Again, it is instructive to cast these assertions in terms of what is being denied, an exercise I leave to the reader. It is also noteworthy that statements such as the one to the effect that a clear statement of desired performance facilitates performance evaluation border on the tautologous.

*The Visual Literacy Movement:*

*Basing a Research Program on Metaphor*

Neither is Holmberg unique in respect to his elevating metaphor to the status of theory. Cassidy and Knowlton (1983) set out, in an analytical work, to assess the basis of an entire influential research program or "movement" called the Visual Literacy movement (henceforth, VL). The purpose of a research program is to test and elaborate a theoretical framework. However, Cassidy and Knowlton argue that, within the borders of this movement, the concept of visual literacy functions merely as a metaphor, albeit one that points general directions for inquiry and research. Their stated intention is to assess the value of this metaphor from a scientific and technological standpoint, in terms of its fruitfulness. They conclude that VL, far from being helpful, in fact has been detrimental. In their view it has served to gloss important distinctions and has thereby impeded, rather than facilitated, hypothesis generation.

Their analysis is principally a philosophical or conceptual one which targets the

incoherence of the concept of visual learning conveyed by the metaphor. But, as I shall explain, there are in fact three prongs to their attack against VL.

They begin by noting the various purposes of metaphor. A metaphor can serve merely as a symbol that unites otherwise isolated groups and individuals who may share some common interests or concerns. This is quite unproblematic. However, metaphor can also serve as a heuristic, or as a concrete model that informs and directs the investigation of something that is highly abstract and not well known. In this role metaphors, unlike theories, are neither true nor false. Rather they are to be judged as more or less fruitful, more or less capable of stimulating inquiry in useful directions.

In the case of VL, the idea seems to be that if it is important to teach the reading and writing of language, then (by analogy or "metaphorical extension") it is also important to teach the "reading" and "writing" of visual iconic elements such as pictures. The term visual, as it modifies literacy, requires some explanation. Writing, of course, is also perceived visually, so presumably the point of the adjective "visual" in "visual literacy" is to mark the contrast between signs that are arbitrary, conventional or nonrepresentational (as in a system of written language) and those that are "nonarbitrary, iconic or pictorial".

The interesting thing about the VL metaphor is that it is somewhat counter intuitive. In the early days of the audiovisual movement, instructional pictures were considered valuable for their communicative transparency, for the ease with which they may be understood by learners. The VL metaphor reverses this viewpoint, positing that instruction is required if a learner is to acquire the skills necessary to understand pictures effectively.

There are, as I remarked, three arguments against VL advanced by the authors. The first two are based on an analysis of the characteristics of language, those very characteristics in virtue of which it makes sense to speak of literacy skills. If the understanding of iconic elements does necessitate the acquisition of literacy skills, it must be because they have the features of a language. So the first question to be asked and answered is this: what are the basic characteristics of language, and do pictures or graphics *per se* have these, or analogous, features?

In the first place, the authors point out, spoken natural languages feature phonological and morphological subsystems. There are in each language a finite number of minimal distinguishable elements called phonemes, which can be combined according to rules to form morphemes, the smallest units that serve a syntactic function. The creation of a written language involves a process of mapping phonemes and morphemes onto arbitrary signs.

Counterparts to the phonological and morphological subsystems do not exist for iconic representation in general. So there is a very fundamental, very significant distinction between iconic and verbal sign systems. The former comprises a finite set of elements that can be arranged according to a finite set of rules. In the latter case, there are no such established or identifiable elements or rules. Instead there is an indeterminate number of elements that may be combined in virtually any manner. Without the counterpart to the rules and elements of the phonological and morphological levels of linguistic systems, the "iconic system allows the possibility that any and all orthographic variance may be potentially meaningful" (p. 70).

These considerations alone would seem to provide an argument strong enough to topple the VL metaphor. If there are no counterparts to the morphological and phonological subsystems, and these are a necessary condition for the existence of a syntactical system which is an essential aspect of a language, then iconic visual representations cannot constitute a language, and literacy skills are not involved in their interpretation.

However, not content to leave things at that, Cassidy and Knowlton develop a second line of argumentation. They proceed to investigate whether or not there could still be meaningful iconic elements; whether, for example, certain colours could denote meanings. They cite from Dondis' (1973) work entitled *A Primer of Visual Literacy* some passages which make the claim that there is indeed such a thing as visual syntax.

The arguments which the authors muster to refute this claim are not original -- they are owed to Goodman's (1968) work entitled *Languages of Art* -- but they are well chosen to do the job. Goodman identifies two necessary syntactical conditions that must be present in order for a symbol system to qualify as notational. A notational system is one which can be notated, thereby permitting consistent, unambiguous communication among those who share the system

One condition, called "character indifference", is defined by Goodman in the following way: "Two marks are character indifferent if each is an inscription (i.e., belongs to some character) and neither one belongs to any character the other does not" (cited in Cassidy & Knowlton, 1983, p. 71). For example, the letters of our alphabet satisfy this condition. The marks *R* and *R* both belong to the character *R* and neither belongs to another character that the other does not.

Cassidy and Knowlton argue, using the example of a splotch of colour, that iconic elements lack this feature. Start with a blot of red paint on a piece of white paper. Now imagine mixing some yellow into a portion of this patch of colour, resulting in a "minute but perceptible" difference in hue. Are the resulting sections character indifferent, or not? They cannot be said to be character indifferent. Both marks might be said to belong to the mark *red*, but just as easily the section adulterated with the small portion of yellow might be said to belong to the character *orange*. Now, obviously a scheme could be concocted to assign marks unambiguously to characters based on this kind of variation. But the point is that such a scheme does not already exist; iconic signs are not intrinsically notational in nature.

The second condition specified by Goodman is that of "finite differentiation": "For every two characters  $K$  and  $K^1$  and every mark  $m$  that does not actually belong to both, determination either that  $m$  does not belong to  $K$  or that  $m$  does belong to  $K^1$  is theoretically possible" (Goodman, 1968, pp. 135-136). Again, this condition fails in the case of the visual iconic. Because iconic elements are inherently continuous (or "analogous") in nature any minor variation on a mark may be significant (or not). To illustrate the point, Cassidy and Knowlton utilize three curved marks which are similar but vary in both size and width of line. It is impossible to say whether the third mark belongs to the second (it shares the same width) or the first (it is closer in size), or neither (perhaps both types of orthographic variation are significant). We could create a scheme to handle this; but, again, the point is that such a scheme does not exist inherently in, or is not internal to, iconic representations.

Cassidy and Knowlton carry things yet a further step by including a discussion debunking studies carried out in the VL tradition that purport to show that there are "literacy" skills involved in the recognition and interpretation of pictorial elements. This is the third

argument alluded to earlier. They argue persuasively that many of the arguments that have been advanced for the teaching of VL rest on a conflation of maturation and learning. I will not go into this lengthy analysis, since the preceding two sets of arguments which I have just recounted seem quite sufficient. The crux of the matter, as they conclude, is: "How can one teach a competency in which there are no identifiable elements nor identifiable rules?" (p. 72).

In sum, we have the evidence for an entire research tradition that is not informed or shaped by any theory but rather only by a metaphor, and one which is arguably conceptually incoherent at that.

### *Theories of Instruction and Instructional Design*

Thus far, I have addressed a number of confusions concerning the nature of theory in educational technology. I have considered the distinction drawn by Holmberg between predictive and descriptive theory, and I have shown how metaphor can be elevated to the status of theory -- explicitly in the case of Holmberg's paper, implicitly perhaps in the case of the research program of VL.

An area possibly even more central to educational technology -- instructional theory and instructional design -- provides illustrations of another set of confusions and misconceptions regarding theory. Richey's (1986) substantial work entitled *The Theoretical and Conceptual Bases of Instructional Design* will serve as the focus of our discussion this time. Richey's purpose is to set out the theoretical and conceptual bases of instructional design (henceforth, ID) in order to make a case for ID as an autonomous discipline. In the course of her argumentation she identifies the major current sources of theoretical underpinnings and then presents a proposal for a research program, if you like, that is intended to integrate these diverse foundations -- which include components from communication theory, general systems theory, learning theory, and the existing procedural models of ID -- into one highly consolidated, comprehensive theory of ID. Along the way she introduces fairly standard definitions of "theory" and "discipline", although she defines the latter as an area of study "whose status warrants the development of a separate and distinct theory," (p. 9) and in this regard may have things a bit backward. It is not the importance or significance of an area of study or a profession which



confers the status of a discipline; rather it is the requirement or possibility of a distinct body of theory and a unique methodology for that region of inquiry that confers the title. Richey actually vacillates somewhat in her assessment of ID as an area of study, generally referring to it as a discipline, but on at least one occasion mentioning it as merely a potential or "emerging discipline" (p. 10).

From the outset, however, she qualifies the idea of a theory of ID with a very significant restriction: "The design of instructional materials and programs is not an activity which is rule-dependent to the degree that two persons could tackle the same problem and arrive at an identical solution" (p. 10). Depending on how narrowly the phrase "identical solution" is interpreted, this may be a very telling rider. If the implication is that intuition plays a definitive role in ID processes, then the way is open to argue that there is no real technology or prescriptive theory of ID.

In fact, such an argument might be quite tenable. Current state-of-the-art practice may suggest that the process is really largely an art or craft which, though it may be organized or structured as a process through procedural models of the kind with which we are familiar, cannot be treated as a science. This is all the more plausible when one considers the central role of formative evaluation (the ongoing appraisal of instructional materials under development for purposes of revision and improvement) in ID models. The real guarantee of success lies in the quality and extent of evaluation rather than any particular set of prescriptive design rules which might be advanced. Thus, ID differs enormously from other design oriented areas where it is uncontroversial to say that a technology exists, such as the fields of engineering.

Richey begins by establishing the scope of ID. Her definition includes the specifications for supervision and maintenance of instruction, as well as all the stages of "micro-design" and "macro-design" (i.e., of materials and programs). The scope of this definition is more expansive than Reigeluth's (1983), which restricts the specifications generated in ID to the development and combining of instructional materials; but less inclusive than Brigg's (1977), which includes materials development. She then proceeds to explain the role of theory in ID. Problem-solving in ID is said to be dependent on an understanding of the processes and relations operating in an

instructional context and what effects are produced by manipulating these processes and relations -- and this understanding is clarified by theory which is "descriptive, explanatory and prescriptive" (p. 12).

ID in the macro-sense is also compared with curriculum theory, and Richey's conclusion is that there is little difference except in terms of orientation: the bases of ID suggest a process orientation, while the primary focus of curriculum theory is a concern with subject matter. There is also some discussion of how ID as a pragmatic and goal-oriented activity meets with the requirements of a profession as put forward by Herbert Simon in his *Sciences of the Artificial* (cited in Richey, 1986, p. 14). Note, though, that the concepts of a profession and a discipline are not necessarily coextensive.

#### *The Relations Among Theories and Models*

More significant for our purposes, however, is Richey's discussion of the nature of theories and models and the relations among them, and her ideas concerning how an integrated theory of ID can be developed. Let us first of all address the issue of models and theories. To begin with, there is some discussion of deductive versus inductive theories that parallels ground already covered in conjunction with Holmberg (Richey, 1986, pp. 13-25). Models are then characterized as micromorphs (visual) or paramorphs (conceptual, procedural or mathematical), with subsequent emphasis on paramorphs. A brief explanation of each class of paramorphic model and its relation to theory is in order:

1. Conceptual models: These are described as the closest thing to theories. They specify, and fully define, all the relevant components. However, they are not as fully explanatory as theories in virtue of the fact they do not contain "clear statements of law or propositions which are supported by quantities of systematically collected data" (p. 17).

2. Procedural models: These are prescriptive in nature. Richey says that, ideally, specific procedural models would be grounded in a confirmed theory rather than based totally on practitioners' experience.

3. Mathematical models: These play two roles, according to Richey. Typically they quantify the relationships which are conveyed in a theory. However, they can also serve to

furnish a more tentative, hypothetical model of relations. In the latter role they do service in the task of theory construction, rather than in theory testing or theory translation.

In general, Richey concludes, models are employed in theorizing in order to organize and integrate knowledge accumulated from a variety of sources. In this capacity they serve to stimulate the generation of hypotheses and the development of theory. However, they also serve as a device for "translating theories into concrete terms suitable for application to practice or theory testing" (p. 17). So according to the author, models are given to theories, to render them testable (this may involve providing a quantitative model) or to render them useful to the practitioner. She does not draw attention to the fact, however, that these two requirements are quite different. To mathematicize a theory is one thing, to create an applied theory or technology from a basic theory is quite another. The latter may well require a radical shift in semantics and terminology. (I will return to this point when I discuss the distinction between theories of instruction and of learning). And then, again, models (simplified ones, apparently) are also said to be something that precede the development of theories and their final, necessarily more highly specified, models.

Already, it seems, the concept of a model and its relation to theory is blurred in Richey's account. It would be preferable, surely, to differentiate these roles and discursive objects with appropriate terminology. For example, the simplified "models" used in the initial stages of theory development might well be simply metaphors or analogies, and it would be useful to call them such. This would also distinguish them from other objects which play a role in the development of theories and which Richey is also wont to call models, such as taxonomies.

The problems with Richey's account of models and theory surface further in the latter portion of the text where different species of the conceptual and procedural categories of model are identified, and where specific models in the literature of the field are categorized. For example, the categories of conceptual models cut across the basic distinctions she has already made by including classes of mathematical formulations and visualizations. The remaining subclasses of conceptual models are "narrative descriptions" and "taxonomies". We are entitled to ask how a taxonomy satisfies the definition of conceptual model reported earlier. A taxonomy

can tell us how to carve up the world in our descriptions, but it will not tell us what are the contingent relationships among the different elements subsumed in the taxonomic scheme.

With regard to the classification of existing models, in the early going Merrill's component display theory is classified as a conceptual model (p. 24); however, in a later chapter it appears as an instantiation of a procedural model (subclass "general presentation model"). This is not entirely surprising. It can be attributed to some extent to the eclectic nature of the model in question, a model which includes taxonomies (classifications of learning outcomes and of presentation forms) and prescriptions based on matching presentation forms and performance levels. But it is also clear that Richey is somewhat ambivalent about the role of procedural models in terms of the development of theory. She writes at one point that procedural models "do not play a clear role in the process of theorizing," but that rather "they reflect current and proposed practice . . . identify steps, not relationships among variables" and therefore they neither explain nor describe events (p. 94). And at an earlier point she acknowledges that there is little connection between the conceptual models of ID and the procedural models developed and utilized by practitioners (p. 20). However, she then proceeds to argue that the current role of procedural models, until such time as there exist well grounded theories to translate into procedural models, is to serve as "a source of knowledge upon which theories can be constructed" (p. 94).

#### *Instructional Theory, Learning Theory and ID Models*

There are other questionable aspects of Richey's synthesis of ID knowledge bases and her proposal for developing an integrated, comprehensive theory. One of the largest difficulties is that she blurs important distinctions between things that are, conceptually, quite significantly different in nature. The distinction between a theory of instruction (henceforth, IT) and a theory of learning (henceforth, LT), for example, is a fundamental one which should be respected. Reigeluth rightly stresses that educators should understand that IT must include reference to specific methods of instruction. Much of what is currently referred to as IT is actually LT (Reigeluth, 1983, p. 23). Landa, too, insists on the distinction, explicitly rejecting the notion that IT should be construed merely as learning theory that would have definite implications for

methods of instruction and for practice. IT focuses on methods of instruction, on "relationships between teacher's actions (or instructional variables) and learner's psychological or behavioural processes." LT, on the other hand, focuses on learning processes. It deals with "relations between learner's actions and learner's psychological or behavioural processes" (Landa, 1983, p. 62).

Richey, however, does not seem overly determined to maintain these distinctions. She admits that there is nothing in existence which amounts to a formal theory of instruction (p. 73) and also that there is controversy over what shape IT should take (pp. 73-74). However, in a section entitled "Procedural Models of ID Based Upon Learning and Instructional Theory," she claims that there is no major distinction to be made between the influence of learning theory and conceptual models of instruction, since "to a great extent, the instructional models have been derived from various learning theories" (pp. 116-117). The word "derived" is certainly too strong here. One cannot derive, in any strict sense of the term, a theory that prescribes methods of instruction from one which essentially does not refer to methods in its vocabulary. Richey seems herself to overestimate what she refers to as a "a direct logical development of thought from learning to instruction to ID -- (a development) so logical that many have blurred the boundaries between them" (p. 57).

#### *Richey's Prescriptions for ID Theory Development*

At this point, we should shift our focus to examine Richey's plan for synthesizing all these diverse sources of knowledge -- from general systems theory, communications theory, learning theory, instructional theory (or, as Richey more correctly labels them, conceptual models of instruction) and existing, predominately procedural, models of ID -- into one coherent theory of ID. This plan, I shall argue, (a) ignores crucial distinctions mentioned above, and (b) exhibits common misunderstandings concerning how theories are constructed by a process of successive low-level empirical generalizations.

As mentioned earlier, there are three components to Richey's proposal for developing ID theory. The first is to come up with a comprehensive, unifying conceptual model of ID. The second is to explain how a theory can be developed and tested from this model. The third is to

explain how the theory can then be translated into models for practical applications.

In order to come up with the conceptual model Richey undertakes an exhaustive cataloguing and classification of related models and theories from the sources referred to above. A partial list would include contributions from Briggs, Gagne, Merrill, Reigeluth, Dale, Schramm, Snow, Bruner, Carrol, Bloom, Krathwohl, Levie and Dickie, Grayson, Ausubel, Marsh, Wright and Payette. It is important to note that some of these contributions represent centrifugal tendencies in conjecturing about the nature of instructional phenomena. For example, Snow argues for localized theories of instruction, while Reigeluth believes in the possibility of a more global, general theory. Still, Richey proceeds to categorize all the variables addressed by these different items. She comes up with a first level classification scheme comprising four clusters of variables relating to the learner, and the content, environment and delivery of instruction. She then proposes that the model can be completed via a simple linear regression equation that will capture the general relations among these clusters and their components: 1. Achievement =  $b_lL + b_cC + b_eE + b_dD + e$  (p. 133).

Or, more correctly (including all possible interactions):

2. Achievement =  $b_lL + b_cC + b_eE + b_{lc}LC + b_{le}LCE + b_dD + e$  (p. 201).

These clusters, Richey admits, are not mutually exclusive. Also, each one breaks down into second, third and fourth level classificatory schemes. For example, the cluster called "learner" can be broken down into demographics, capacity, competence, and attitudes. Capacity is then subdivided into: intelligence, cognitive development, and physiological development. And, finally, any of these can in turn be subjected to further analysis. For instance, cognitive development could be broken down in terms of a taxonomy such as Bloom's (1956) hierarchy of cognitive behaviours.

It is striking that one would contemplate beginning with such a soup. The various taxonomies, theories and models addressed in Richey's review are sufficiently diverse in origin and import that it is not obvious that they must be systematizable. Moreover, many of them may not even have any validity in themselves; for example, Richey includes items which are no longer widely accepted, such as Ausubel's Subsumption Theory (Ausubel, 1963; 1968). There is no

doubt a need for a common conceptual framework from which to begin theorizing, but it is far from clear that the way to achieve such a framework is by including all the variables mentioned in all the existing theories and taxonomies. Such a list can hardly be called a conceptual model. Nor can the simple regression equation she recommends constitute anything approaching an interesting account of the relations among the different items.

Any integrative conceptual model must be one which suggests just what is the basic set of fundamental explanatory variables that needs to be explored. It is not at all evident that Richey's proposal advances the current state of the art in this regard. In effect she provides a structured list of variables that have already been considered in past research. Even if a stable regression equation could be discovered for these clusters of variables, it is not evident how that would facilitate the formulation of a theory. Knowing how much variance in a system is attributable to different variables simply does not tell us what are the fundamental causal principles underlying this variance.

There are methodological problems to go along with the analytical ones to which I have just drawn attention. The regression equation cannot be used as a structure to synthesize and consolidate past research. We have no way of separating out sources of error and assessing the relative contributions of different interactions when we address research that has already been conducted and that by design may not include provisions for all the various possible interactions which are represented in the equation.

There remains, of course, the possibility of utilizing the regression equation as a framework for a new research program within the discipline of instructional design. Richey suggests that studies can be conceptualized by:

systematically altering key variables to compare resulting regression coefficients . . . If researchers operate with a common set of variables, using consistent analyses, differing beta weights can be highlighted, thereby showing the effects of alternative contexts. (p 207)

The ultimate goal is then to integrate results from a variety of studies that investigate different facets of the four clusters (environment, learner, content, and instructional delivery)

which Richey has posited. The hope is that consistent use of the regression model across these different studies will allow us to determine whether or not the structure of the relationships among the basic variables is stable. What is anticipated is a general theory that subsumes many types of system. "If comparable beta weights were produced (in different contexts), we could propose such generalizations" (p. 207).

It is far from self-evident, though, that such comparable weights would result, since the error term would likely fluctuate widely among studies, reflecting the unfathomed variability in system effects and interactions across these different settings. Even if it did not, there remains the possibility that the error term associated with variations in the system that are not captured in Richey's taxonomies would remain large. The problem then would be to determine what variables or underlying structure might be reflected in that error, as these variables and structure should be incorporated in any adequate theory.

There is also a difficulty associated with the requirement that we employ equivalent measures across these different contexts. The cost of developing and validating equivalent measures appropriate for such a range of contexts as Richey envisages would be prohibitive, even if it were feasible in principle. Yet without equivalent measures, any comparison of the weights associated with different variables across the studies would be meaningless. So, too, would the analysis of data pooled from the studies. Yet in the absence of these conditions (the assumption of equivalent measures and the appearance of comparable beta weights across studies conducted in divergent contexts), researchers are in no better position than they are presently. They are simply confronted once more with the tasks of sorting out sources of variability, identifying the basic variables and their interactions, developing reliable and valid measures of these constructs, and grappling with the question "To what can we generalize?".

In short, a regression equation, even a stable one, does not constitute a theory, and the path from a stable formula to a working explanatory model is not a straightforward one. Neither does a list, or even a set of taxonomies, amount to an integrative conceptual model that can serve as the trappings of a theory.



### *Summary*

To summarize, Richey distinguishes between what she refers to as "formal" theories and models of various kinds, including the conceptual, procedural and mathematical varieties. The role of a theory is identified as the integration and systematization of our knowledge, coupled with the provision of explanations and predictions. This is a fairly orthodox account which is palatable enough. However, problems surface when Richey begins to explicate the various categories of models and their roles in theory development. As mentioned above, the basic distinctions among conceptual, procedural and mathematical models are eroded by the inclusion of mathematical models as a species of the conceptual category. Moreover, the characterizations of these various objects is not completely consistent. At one point the author insists that a conceptual model should identify all relevant variables. However, she then proceeds to label Dale's Cone of Experience as a conceptual model. It is in fact merely a taxonomy which classifies media according to a single attribute: closeness to reality. Of course, one can reply that a distinction has to be made between a good conceptual model (which will be as complete and as many faceted as required) and a poor one (which may address only a small subset of the variables which need to be included). But there remains the difficulty that Richey is allowing a taxonomy to qualify as an integrative conceptual model. A mere list does not deserve to be called a conceptual model if, as Richey holds, conceptual models are the stuff from which theories are spun. As remarked earlier, an integrative conceptual model should go at least a step further; it must convey suggestions regarding the basic relations and mechanisms underlying the phenomenon of interest.

Commendably, Richey insists at the outset on a distinction between models and theories. Without this distinction it is difficult to even discuss the processes of theory development and elaboration and theory testing. It is not a distinction that has been universally respected in the field, however. Dick (1981), for example, holds the opposite view: "It may be argued that generic ID models represent the theory of ID. The theory includes a description of a series of steps which, when executed in sequence result in predictable learning outcomes" (In Richey, p. 24).

Richey's position seems to be more defensible than Dick's stance. Even if the sets of procedures Dick alludes to are elaborated in sufficient detail to constitute models, and even if we grant for the sake of argument that these models have predictive power, they still fall short of the requirements of theory in the accepted sense of that term. Specifically, procedural ID models do not generally rate very highly in terms of explanatory power and this quality is one of the defining characteristics of theory.

Unfortunately, once having made this distinction, Richey proceeds to come very close to violating it herself. On several occasions, as noted previously, she appears sympathetic to the position that the significance of distinctions among conceptual models of instructional design, theories of instruction and theories of learning, may be minimized. Ignoring the boundaries among these objects is especially harmful if one's principal task is to examine the potential for the development of a theory of instructional design. Such a project must address, at its core, the interplay, the historical and logical relations, among these very objects.

It is in fact in the area of proposing a strategy for theory development that Richey is most open to criticism and where her subscription to naive conceptions of theory and the processes of theory development is more thoroughly exposed. The distinction she introduces between inductive and deductive theories in chapter one of her book is essentially a spurious one which reflects standard misconceptions that have achieved a certain currency. I noted the same distinction in my discussion of Holmberg's work. More importantly, her proposals for utilizing the framework of a list of previously identified variables (compiled from work that exhibits, in some areas, incompatible or centrifugal tendencies) are untenable. They can be criticized for their apparent naivety and simplicity. But the real objection is that they seem to confuse the methodology of hypothesis testing, where regression equations may be featured usefully, with the complex processes and dynamics involved in inventing theories. Even a stable, highly generalizable regression equation does not deserve to be called a theory. Nor is the progression from such a low level empirical generalization to a full blown explanatory theoretical framework an obvious one. In brief, inherent in Richey's proposal for the regression equation framework is a confusion of product and process in science.

### *Further Examples*

I have tried to illustrate a number of confusions and misconceptions regarding the nature of theory which seem to be widespread in the literature. One major problem is the confusion of procedures for testing theories with the process of theory elaboration and perhaps even with theory itself. Another is the elevation of very simple forms of discourse -- including virtual truisms and common sense generalizations as well as taxonomies, lists of variables, metaphors and analogies -- to the status of theory. Finally there is some confusion over terminology as regards the useful distinction between theory and model.

The examples presented were chosen on the basis of several criteria. They are published in reputable, "mainstream" educational technology periodicals. They concern areas of central concern to the field such as mediated learning, distance education and instructional design, and visual learning. And, finally, they deal quite explicitly with the conceptualization of theory and the task of theory construction.

It is, admittedly, a limited sample. However, it is not difficult to find other examples to substantiate the thesis put forward in this chapter. The literature of distance education, for example, features many articles by leading figures in that field who espouse the view that distance education should be regarded as a separate discipline which requires its own unique theories of instruction, communication and learning (Baath, 1981; Holmberg, 1983; 1986; Keegan, 1980; Perraton, 1981; 1987; Tight, 1988). These articles seem to betray a lack of understanding of the nature of theory. It is true that from a practical standpoint distance education schemes frequently pose unique problems, based on the peculiar administrative and logistical constraints they may involve. But what seems missing from this literature is any careful attempt to demonstrate that, from the point of view of learning and instruction, the phenomena of distance education are different in kind from their counterparts in conventional settings, and thus require the development of separate theories. Indeed, efforts to define distance education in ways that distinguish it from conventional schemes seem inevitably to come down to the physical separation of learner and tutor. It cannot be assumed that this factor entails the requirement for new theories. Distance education schemes may simply represent different initial

conditions that fall within the parameters of more general theories. Indeed, on the instructional side it has been remarked that the literature of distance education has paid scant attention to the literature of instructional theory and the relevant empirical research concerning instructional variables and strategies (Bernard, Naidu & Amundsen, 1990). Any conclusions regarding the need for theories specific to distance education would therefore seem premature.

Further examples of simplistic hypotheses cast in the role of theory can be gleaned from the area of visual learning. In addition to the VL metaphor discussed above, there are a group of hypotheses and taxonomies advanced from the late 1940's through the early 1960's which collectively have been referred to as realism theory. These include Morris' (1946) iconicity theory, Dale's (1946) cone of experience, Carpenter's (1953) sign-similarity hypothesis, Gibson's (1954) projective-conventional continuum, Knowlton's (1964) transparency-opacity continuum, and Osgood's (1953) more detachable-less detachable continuum. The common factor among these elements comprising realism theory is the thesis that the probability materials will facilitate learning is linked to the degree of realism which characterizes them: the greater the degree of realism, the greater the likelihood that they will promote learning. Dale's cone, for example, orders experience from the abstract to the concrete, with the assumption that more lifelike experiences provide for more effective instruction.

Realism theory influenced educational technology theory and practice, particularly through the period of the field's development in the 1950's and 1960's, when it was closely identified with the audio-visual movement. An article by Finn (1953) concerning professionalization in the field published in the journal *AV Communications Review* provides a good illustration of how the audio-visual movement was informed by this thinking. Despite the lack of empirical research supporting realism theory (cf. Travers (1964) for an account of the lack of confirmation and of contrary evidence from physiology and the psychology of perception), it continues to exert some influence. Devices such as Dale's cone are still mentioned in the literature of instructional systems design.

The assumptions underlying realism theory are also still tacitly influencing the field in another area: the question of the importance of "fidelity" for learning or transfer of training in

training system design incorporating simulation (Hayes & Singer, 1989). In practice there seems to be a general belief that the greater the correspondence between the training and operational environments, the more likely is successful transfer to occur. Yet, again, this is not supported by empirical evidence.

The point, of course, is not so much that realism theory is apparently false. It is not even that it has no logical basis; the roots of realism theory can be traced back to associationist frameworks in epistemology and in learning theory. Rather, the problem is that it is misleading to refer to an informal collection of hypotheses and classification schemes, which are loosely related by underlying assumptions, as "theory". Realism theory is simply not theory at all in the strict sense of the term. Despite the fact realism theory is now widely considered to be false by researchers, it is still frequently afforded the status of theory.

One problem with referring to this amalgamation of elements as a theory is that the term implies a certain degree of integration which is not present in realism theory. For example, while one of the basic tenets of realism theory is that a greater degree of realism in instructional graphics enhances learning, there is also a recurrent theme which states that multi-channel instructional communication enhances learning. Logically these two hypotheses appear quite distinct within the bounds of realism theory. It appears then, that it is logically possible to refute certain of the components of realism theory without implications for other basic hypotheses. This degree of independence among constituent propositions is not characteristic of theory and is not desirable in an explanatory framework. The danger inherent in allowing the title of "theory" to be applied to such loosely associated conjectures is that they then serve to condition, in some large part, our understanding of theory in the field. The limitations of this understanding then constrain our efforts to further develop our knowledge. More precisely, there is the risk we may become too satisfied with low-level empirical generalizations and fail to probe more deeply for integrative explanatory frameworks.

*Readjusting the Goals of Research:*

*Low Level Generalization versus Deeper Explanation*

Winn (1989), Clark (1989) and others have made much of this last point with respect to

the conceptualization of research and research questions in the field of educational technology. Clark's complaint, which has been echoed by many, is that we tend to limit ourselves to lower level generalizations. The fictitious example he sometimes uses to illustrate his point concerns comprehension, retention and time-on-task in learning from prose materials (as recounted by Winn, 1989). The educational technologist, the parable goes, will often proceed in the following way. A study will show that reading materials twice rather than once improves learning. A subsequent study may show that reading materials three times is even better, though the gain is smaller than what occurs between once and twice. Finally, a further study may then establish that reading materials four times offers no significant gains (or else perhaps the gains are so slight that the law of diminishing returns can be invoked). The perpetrators of these studies may then conclude from the results of this "research program" that it would be best if learners read all materials, say, twice.

The point that needs to be considered is that there may be a theoretical explanation of why reading something twice improves learning, and that such an explanation could potentially suggest a more efficient way of achieving the same gains. In this example, the theoretical notion that might be invoked would perhaps be "elaboration". The learner who reads materials twice learns more because he engages in greater elaborative behaviour. But there are perhaps other, more efficient, ways to increase elaboration: for example, the use of organizers, summaries, and embedded questions might accomplish the same results with less time spent on task overall.

#### *Educational Psychology as Conceptual Analysis*

A final point regarding the elevation of trivial, common sense or, even, tautologous propositions to the stature of theory is in order, also. Egan (1988) has gone far beyond any claims advanced in this chapter concerning this particular issue. Egan makes the claim that all of educational psychology is essentially tautologous, that the laws of human learning and motivation are analytical truths in the last analysis. His argument is not that the goals of educational psychology are trivial, but rather that the methodology employed is the wrong one. If the basic laws of human learning are really analytical in nature, then the most efficient means to establish them is conceptual analysis rather than the empirical methodology of

experimentation and hypothesis testing. Egan concludes that the formal, deductive theorizing of geometry provides a better model for psychology than the empirical sciences.

Among the examples Egan uses to make his case are the following:

1. From among a list compiled by Hilgard (1956) of propositions that allegedly have been established by empirical psychology Egan selects the hypothesis that "brighter people can learn things less bright ones cannot learn" (Egan, p. 72). He points out that there is a conceptual link between brightness and ability to learn, in virtue of which this proposition is true

2. Egan also considers another proposition from Hilgard's list: "a motivated learner acquires what he learns more readily than one who is not motivated" (Egan, p. 72). Again, he points out that what we mean by a "motivated" learner is conceptually related to what we mean by learning more readily. One cannot conceive of the motivated learner without invoking behaviours which instantiate more ready learning.

3. Finally, Egan considers the supposed empirical connection between ease of learning and degree of organization of lists. There is, he argues, a necessary connection between order and learning:

A detailed definition of learning would imply or involve the notions of order; the structure of the human mind, important for what can be learned and how, and what is conceived of and recognized as ordered are not distinct things. (p. 73)

I would not debate the point that empirical psychology betrays numerous elements of an analytical nature. Egan is also not alone in making these claims. He cites the work of Louch and Smedslund who have argued the same point. However, the conclusion that all of educational psychology must be analytical in nature is more radical than his arguments or examples support. It is not so obvious, for instance, that detailed, comprehensive cognitive theories which deal with the structure of memory and the nature of knowledge representation are just compendia of so many analytical truths. The examples Egan chooses are unfortunate ones for empirical psychology, but one can doubt whether, for example, Anderson's (1983) ACT theory or its various applications to specific learning tasks and domains would be amenable to the same analysis. Furthermore, the examples are unfortunate ones not, essentially, because

they are tautologous – since virtually any empirical theory will contain some analytical elements (definitions) -- but rather because they are trivial and because they have been mistaken for significant truths.

There is also the following point to be considered: even analytical truths can sometimes only be established, initially, through empirical means. For example, a whale is by definition a mammal, so the statement that a whale is a mammal is an analytical truth no less than a statement such as "all bachelors are unmarried men." However, it is conceivable that the whale might be mistaken for a fish until such time as a detailed empirical examination of the creature could confirm that it does indeed have the requisite characteristics of a mammal. Thus where questions of classification are concerned ("Is X an instance of Y?"), and this would appear to encompass much of psychology, the fact that an analytical truth obtains in a certain context may only be determinable by empirical means. It may well be that some object X is an example of Y, and hence has the characteristic P (a defining characteristic of Y) by definition; but establishing the fact that X is a case of Y may require investigation of X by empirical means. This consideration alone would suffice to distinguish psychology from purely formal disciplines which proceed entirely from definitions and which require only consistency of their theories. It therefore serves to undermine the analogy Egan tries to develop between psychology and geometry.

### *Conclusion*

The argument of Chapter 3 was that that educational technology research and theory development is limited by the widespread adherence to faulty conceptions of theory evident in the field and by erroneous views concerning the processes by which theories are developed and refined. The influence of operationalism and the weakness of procedures for testing theories and hypotheses were emphasized, in general terms. In the present chapter I have tried to concretize and refine these criticisms. Several naive or simplistic conceptions of theory and theory development have been illustrated using examples drawn from the literature of mediated learning or distance education, instructional design, "visual" learning and "realism" theory -- all major areas of research within the field of educational technology. The misconceptions that



have been exhibited include metaphor as theory, tautology as theory, taxonomy as theory, and low-level generalization as theory. The misconceptions concerning the development and refinement of theory include a naïve inductivist view which holds that theories are built up piecemeal from an aggregation of low-level empirical generalizations, and the confusion of process and product (the formulation of theory as opposed to the testing of theory) evident in Richey's proposal for theory development in instructional design.

The basic point which needs to be stressed is that we can never hope to progress to the development of sophisticated theories that will explain and predict some of the complex instructional phenomena which interest the field unless we first understand what a theory is. However, it needs to be emphasized, too, that the mistakes discussed above are naïve but not "silly". There certainly is a relationship between metaphor and the models provided for sophisticated theory. Every theory invariably begins life as a metaphor. But it is important to realize that a well-developed theory involves more than a metaphor. It requires a model which is specified in sufficient detail to allow exact predictions and hence rigorous tests. It is also true that the formulation of low-level empirical generalizations typically precedes the elaboration of deep explanatory theories. But no number of such generalizations constitutes a deep, consolidated theory, and the route from a heap of generalizations to an elegant, systematizing conceptual framework typically involves a leap requiring great insight and creativity.

A first step in reforming our conception of theory would be to adhere strictly to distinctions marked by the terms metaphor, taxonomy, theory and model. The term model, in particular, is notoriously ambiguous and is frequently used as a synonym for theory. Andrews and Goodson (1980), for example, constantly interchange the two terms in their comprehensive review of instructional design models. The formal distinction between a theory as a set of syntactic elements or propositions and a model as the interpretation or semantics attached to those linguistic elements is a particularly useful one, and should perhaps be adopted as standard usage. A model is roughly defined as a set-theoretic entity comprising an ordered set of objects together with the operations and relations associated with those objects. (A more thorough definition is provided in Chapter 5.)

## CHAPTER 5

### The Relationship Among Alternative Paradigms

In Chapter 2 I outlined some of the basic precepts of two alternative paradigms which are currently vying for prominence in educational research: the interpretive or qualitative and the objective or scientific. In that chapter I was concerned primarily with characterizing salient aspects of the scientific tradition and assessing its viability. This pursuit led us to specify the basic elements of the qualitative paradigm, since most of the current criticisms of the scientific approach in applied social science fields is made from the perspective of alternative frameworks. These alternatives include critical theory and various species of systems thinking, but the qualitative camp figures most centrally.

My task now is to deal with a rather complex question: what is the relationship among these alternative paradigms, i.e., the qualitative and the quantitative. Until this question has been answered we cannot claim to have assessed the role and potential of scientific inquiry in educational research. It is an important question and especially topical, given that the scientific approach is no longer preeminent in the field of education in general.

The emphasis on the relationship between qualitative and quantitative approaches requires comment. The justification for this focus is twofold. In the first place, as just mentioned, qualitative approaches have acquired an increased following in general educational research, and we are beginning to see, more recently, reflections of this movement in the narrower subfield of educational technology. Secondly, the relationship between qualitative and scientific paradigms is perhaps the most problematic among paradigms. There is considerable debate in the literature concerning whether and how these approaches either conflict or intersect. Moreover, there has also been a strong trend towards evolving hybrid approaches which combine elements of quantitative and qualitative research. Part of my goal in this chapter is to evaluate the case for this sort of composite methodology.

As Smith and Heshusius (1986) have observed, we seem to have moved through three basic phases as regards the field's perception of the relationship among these paradigms: a period of mutual disdain, followed by a fairly short transitional period of relativism in which

different paradigms were viewed as distinct (possibly incompatible) but equally valid, and then finally the current period with its emerging attitude of cooperativism and its methodological trends towards the integration of different approaches.

The general historical tendencies are easy enough to see. Separating out the different precise positions which have been taken regarding the relationship among paradigms, and assessing the arguments which have been made in support of each, is not so easy. It is made all the more difficult by the circumstance that both researchers and philosophers who have addressed this question have generally failed to make the very sharp distinctions among possible viewpoints that are required, so that it is not always even clear exactly what position is being defended or attacked.

My line of approach will be as follows. I will first identify and characterize, exhaustively, the different positions which may be taken. I will then proceed to elaborate and evaluate the principal arguments for and against these positions. In the last analysis, the one which stands up best to scrutiny is the one which I have labelled "incompatibilism". This is a conclusion which was anticipated, to some degree, by the arguments and portrayals of Chapter 2. Incompatibilism, in turn, undermines the case for hybridization of methodology. If incompatibilism is the correct interpretation of the relation between qualitative and scientific paradigms, then such a strategy at best leads to a weakening of scientific methodology, where what is called for based on the arguments of the last chapters is a strengthening of certain areas.<sup>9</sup>

### *The Alternatives*

#### *Incompatibilism*

This is the thesis which asserts that the alternative paradigms involve precepts which are mutually contradictory. It follows that they cannot rationally or fruitfully be combined in the hybrid approaches alluded to above.

The *prima facie* case for incompatibilism is strong, given that the naturalistic and scientific paradigms do involve fundamental assumptions that appear clearly contradictory. As I

have argued, these assumptions involve, minimally, assertions concerning the nature of explanation, the possibility of generalizations and predictions, the meaning of truth, the meaning of internal validity, the nature of social reality, and the relationship of the inquirer to the phenomena investigated. However, the case for incompatibilism must also be constructed largely in negative terms, by refuting the alternative positions. It must be conceded that the weight of thinking among practicing researchers, and also among philosophers of social science, appears to be in favour of the other options named below.

### *Orthogonality*

According to this position, the most accurate description of the relation among paradigms is that they are irrelevant to one another. The basis for this claim is the idea that each paradigm is appropriate to a specific range of purposes or a specific class of inquiries and that, in this respect, they are mutually exclusive.

In short, since the different paradigms address different types of questions, they cannot be viewed as either conflicting ("incompatibilism"), or mutually consistent ("compatibilism", below). Nor can the choice between them be seen as an arbitrary one ("relativism", below). Given a certain range of purposes or interests, one is constrained to choose the relevant paradigm. The different basic paradigms can thus be said to be complementary, but only in the sense that each addresses problems of a type not directly touched by the others.

### *Compatibilism*

This is the view that the different paradigms are compatible with one another. Hence, on this view, the results of the different methods of inquiry can be compared and, where they agree, can be regarded as mutually supportive or reinforcing from an evidentiary point of view. The paradigms can thus be regarded as "complementary", though in a sense different from that mentioned in the context of the thesis of orthogonality.

Compatibilism thus carries two important methodological implications. First, it establishes at least the logical possibility of hybridization of methodologies, though alone it does not go so far as to furnish a sufficient argument for blending the elements of various modes of

inquiry. In order to construct such an argument one would have to establish the additional premise that some amalgamation of approaches would furnish a method of inquiry more powerful than any unitary perspective. Secondly, it forms the basis for another methodological strategy, namely, triangulation (Mathison, 1988).

The relationship of compatibilism to triangulation is much tighter than with hybridization. The idea underlying triangulation is that each paradigm has its own inherent bias. By bringing a variety of approaches to bear on the same investigation, one achieves a sampling of sorts, in which the biases of different methods supposedly tend to cancel one another out. As these different methods provide a variety of evidence in support of a specific conclusion or set of conclusions, they thus bring us, by a process of convergence, to the "truth". The more difficult question raised by the concept of triangulation, of course, is where one is left when convergence does not occur, and what this implies for the presumed compatibility of paradigms. The interpretivists' caution that there are multiple social realities and that convergence should not be anticipated, even within qualitative studies, also raises questions about the logic underlying this form of triangulation.

### *Relativism*

According to the thesis of relativism, there is no rational means for selecting among paradigms; they are all equally valid. Relativism is based on arguments stemming from either of two philosophical positions.

The first of these is associated with Kuhn's conceptualization of paradigms, and is called incommensurability. To say that paradigms are incommensurable is to say that there is simply no way to compare the contents of different paradigms, that the logical relations of consistency, contrariness, contradictoriness or entailment do not apply across paradigms. The incommensurability thesis, in turn, rests on what I shall call the principle of "meaning-variance": the notion that there is no neutral, paradigm- or theory-independent language. The relationship between the two is easy to see. If the results of different theories or different research traditions are not expressed in, nor translatable into, a common language, then there can be no basis for

asserting any comparisons among them.

The second basis for relativism is a set of philosophical positions which are all concerned with the underdetermination of theory by evidence. These include Quine's peculiar brand of pragmatism, the Duhem thesis and, most importantly, conventionalism. These will all be explained and evaluated in due course. For the moment, I will simply present a brief statement of conventionalism, and a comment on its foundation.

Conventionalism is the doctrine which asserts that there may exist rival, incompatible theories (and hence associated paradigms) which are indistinguishable as regards the statements or predictions they make concerning the observable world. Conventionalism thus rests upon the assumption of a predominately theory-neutral observational vocabulary for science, and locates the incompatibility among theories in their respective, peculiar, theoretical vocabularies. Confronted with a set of theories which satisfy the requirements of conventional alternatives, one has no epistemic basis for choosing among them. Hence one is free, it seems, to select any. We thus arrive at relativism along a path which, interestingly, involves assumptions concerning the observational-theoretical bifurcation that are in direct opposition to those which underlie the route via incommensurability.

### *Compatibilism*

#### *Pragmatism as a Basis for Compatibilism*

The most recent plea for embracing compatibilism has come from Gage (1989). In an article entitled "Paradigm Wars" the author projects twenty years into the future and sketches three alternative scenarios:

- (1). The scientific ("objective-quantitative") approach has been superseded by naturalistic inquiry and critical theory.
- (2). Things are much as they are now (meaning the paradigms are separate and regarded as fundamentally at odds).
- (3). A full rapprochement or accommodation has occurred between the basic paradigms.

In the course of the article Gage clearly identifies (2) with the potential demise of social inquiry, and proceeds to plump for (3) as the most desirable or constructive outcome. Within this third scenario, the role of objective inquiry has been severely restricted. There is no question of the possibility of developing high level theoretical discourse. The role of scientific method is limited to tracking and describing social trends, with an aim to providing information required for policy decisions ("social engineering").

Gage is actually somewhat ambivalent as regards the exact nature of this postulated rapprochement. At one point he suggests that the basis for it is to be orthogonalism: "First, it became apparent that programs of research that had often been regarded as mutually antagonistic were simply concerned with different, but important topics and problems" (p. 7) Yet, on the same page he goes on to suggest that product-process research is an inevitable orientation in his area (teaching), and that interpretive, ethnographic approaches can play a role in this mode of inquiry. This seems a relatively clear statement of compatibilism.

What is interesting about Gage's polemic is the basis on which he entreats us to move to (3). In the last paragraphs he calls for educational researchers to reaffirm their moral responsibilities, to recall the moral and rational bases of their enterprise, namely, the goal of improving education, and in so doing to *invent* the future. We are to carry educational research forward in a direction that is meaningful and fruitful, the direction of (3).

Thus, Gage's line of argument seems to presume that whether the various paradigms are to be regarded as compatible with one another or not is a question requiring a decision, rather than careful conceptual analysis and rational argumentation. A desire for a state of "productive harmony", and a recognition of shared values at a very general level is seen as sufficient to establish compatibilism. Or, at least, it is seen as sufficient when combined with a certain element of pragmatism, as the following passage, related to the scenario of rapprochement, reveals. In this scenario it is supposed that a new generation of researchers, alarmed by the threat to social research posed by interparadigmatic bickering,

began to come to their senses. They understood well enough that scientists should learn from philosophers' analyses of their concepts and methods. But they also understood that the philosopher's of science should accommodate their analyses to what scientists actually did. They began to be influenced more by the old-fashioned pragmatism. They realized the moral and rational foundations of the three paradigms were virtually identical, dedicated to the same ideals of social justice and democracy and the goals of an education that would serve those ideals. (p. 8)

This appears rather woolly-minded. What social inquirers do must be compatible with their assumptions regarding, among other things, the nature of the object of inquiry. If their practices are not consistent with their paradigmatic assumptions, then it is the job of philosophical analysis to bring this to light. And it is the researcher's responsibility, in turn, to either revise his assumptions (which then require justification) or change his practice.

Paradigms of inquiry are epistemological in nature. The fact that proponents of different paradigms may share certain ideals does not make the paradigms any more compatible; no more does the fact that there are often ideological differences among the constituencies of rival paradigms serve to establish their incompatibility as methods of inquiry. Values and ideologies may speak to the motivation to adopt a certain mode of inquiry, but they are essentially irrelevant to assessing the characteristics of these as methods of acquiring knowledge and validating claims to knowledge.

The notion that we should be more "pragmatic", that philosophical analysis should reflect to a greater extent what researchers are actually doing, also requires comment. In the field of educational research it is a viewpoint that has been strongly echoed by others. In response to those who balk at the notion of combining paradigms, Miles and Huberman argue that "epistemological purity doesn't get research done" (1984a, p. 21), while Howe (1988) decries what he regards as the tyranny of the epistemological over the practical: "Why should paradigms determine the kind of work one may do with inquiry . . . The possibility of modifying a paradigm . . . in response to the demands of research . . . seems to go unnoticed" (p. 130).



The difficulties associated with pragmatic epistemology are well documented, especially the central criterial problem. (What works cannot be identified with what is true, at least not straightforwardly so, since true statements may be relatively useless and statements that are patently false may have a very high degree of utility in certain contexts.) But it is not necessary to become enmeshed in a discussion of the merits of pragmatism in order to criticize the stance that has just been presented.

The difficulty has to do with the understanding of the notion of a paradigm. Recall that in Chapter 1 I presented Kuhn's conception of the paradigm as based on concrete instances of successful theorizing:

some accepted examples of actual scientific practice -- examples which include law, theory, application and instrumentation together -- provide models from which spring particular coherent traditions of scientific research. (Kuhn, 1970, p. 62)

Recall, also, that I argued that in the social/behavioural domain, research traditions are not based on such models, simply because conspicuous successes do not abound. Thus, for example, our scientific "paradigm" is a research tradition based on a (rather poor) reconstruction of models from another domain.

Paradigms, understood as reconstructions and, in varying degrees, formalizations and idealizations of, concrete practice can, of course, be criticized if they are too far removed from that of which they are reconstructions. Indeed, this was the very basis on which I criticized the operationalist elements of our present scientific tradition: they are clearly in contradiction with actual, successful, scientific practice. However, for actual practice to refute a paradigm, the models appealed to must have a track record of success. Thus, the argument that hybridization is justified because it is what researchers are doing simply does not go through. The additional premise which is required, stipulating that these new practices are more successful, is simply not available at this point. Yet obviously this premise would be crucial to a justification of these approaches from within a pragmatist's epistemological framework.

*Denial of the Fact-Value Distinction as a Basis for Compatibilism*

In Chapter 2 I considered Howe's arguments concerning the fact-value distinction in the context of their implications for the viability of the scientific approach. Howe (1988) concluded that the apparent choice between what he calls the qualitative-interpretivist and quantitative-objectivist paradigms is just that: merely an apparent one. He argued that what is said to distinguish the two, fundamentally, is the role of values (both the subjects' and the inquirers') in each tradition. The qualitative-interpretivist paradigm insists that the researcher's own values influence his conclusions, and that an understanding of the subjects' values is crucial to understanding and describing their behaviour. The quantitative-objectivist approach, in contradistinction, deliberately excludes consideration of values in both respects, according to Howe.

It is worth remarking that there is no reason in principle why "quantitative-objectivist" research could not incorporate values and intentions as (dependent or independent) variables of interest, with a view to their potential as part of an objective explanatory framework. That quantitative research in educational technology has tended not to address these types of variables is largely a contingent historical circumstance. But I have allowed the assertion that the qualitative and scientific paradigms differ in the role they assign to values in explanations and descriptions of social/behavioural phenomena, and in the legitimation of knowledge claims. So while I do not accept Howe's distinction between the paradigms without qualification I can allow his basic thrust, for the sake of argument.

Howe's contention, then, is that there is no real choice between the two paradigms, because there is really no clear defensible distinction between empirical facts and judgements or statements of valuation.<sup>10</sup> Thus, the profile of a quantitative science which describes the social world in purely descriptive, value neutral terms is alleged to be a mere chimera.

While I have rejected the standard arguments arraigned against the fact-value distinction, I shall note what these arguments, had they been sound, would have entailed for the relationship among paradigms. The implication is that scientific research is just as value-laden as qualitative

research; the major difference is that the influence of values in scientific inquiry is not a conscious one. It is tempting to view this as evidence for compatibilism. But from within the same perspective scientific inquiry may be regarded as flawed: the unconscious or tacit nature of the influence exerted by values may be interpreted as a weakness of the method. Thus, while the argument to the effect that all research is fundamentally, irrevocably value-laden can be used to support compatibilism, as a logical position, it does not serve well to recommend triangulation or hybridization.

#### *Summary of the case for Compatibilism*

We have seen that compatibilism cannot be established on the basis of pragmatism and an appeal to what researchers are actually doing and "what works". We have also seen the failure of the attempt to show that a fundamental distinction between the paradigms, the role assigned to values in inquiry, is untenable because the basic fact-value dichotomy on which it rests is unsupportable. We can also note that even if this distinction were collapsed there would remain significant differences between the interpretivist and objectivist paradigms still to be accounted, although perhaps an argument might be made that ultimately they require the fact-value distinction, and would necessarily fall without it.

Two questions remain to be dealt with, then. The first is this: why does compatibilism seem so intuitively correct to so many researchers, despite the apparent differences among the paradigms? And the second is: what is the harm of trying out compatibilism, of permitting multi-method studies and hybrid methodologies, to see if, in the longer run, they can succeed where more pure disciplinary approaches have so far failed to bear fruit?

Part of the answer to the first question may lie with a certain ambiguity attached to the term "methodology" (Rudner, 1966) that may serve to obscure the significant differences among paradigms. Another factor is that a number of commentators have drawn attention to certain similarities or commonalities between qualitative and quantitative approaches which, while in some cases superficial and in others misconceived, have given the impression that the paradigms can be consolidated. These points will be taken up in some detail in the following

sections, after which I will respond to the second question.

*The ambiguity of the term "methodology".* Methodology may be taken to mean "logic of justification", i.e., the general precepts of an approach to inquiry that concern the passage from evidence to conclusions, what constitutes evidence, how it is obtained, and so forth. In this sense, the principles of a methodology are those principles which provide for the warrantability of assertions; they are the basis of the legitimation of our claims to knowledge. On the other hand, methodology is also sometimes taken to mean merely the specific procedures or techniques employed in research.

In the latter sense, it is quite clear that elements of either the qualitative or quantitative paradigms may be exportable. Qualitative researchers, for example, have long recognized the value of "unobtrusive" measures. More recently quantitative inquiry has shown some greater interest in this kind of tool. Similarly, there is no reason in principle why a qualitative researcher may not find some mathematical tools useful for the process of data reduction or analysis. Thus, it is true, but rather trivial, to say that at the level of tools and techniques, different paradigms may be compatible. The logic of justification simply does not dictate precisely which individual tools will be employed at the level of investigative practices.

Smith and Heshusius (1988), however, have rightly stressed that convergences or similarities at the level of tools and techniques do not entail that the paradigms are compatible or complimentary at the level of logic of justification. And, it is the logic of justification which is most fundamental to characterizing a method of inquiry. Investigators hold that the systematic employment of particular sets of procedures secures for them certain basic qualities of inquiry, such as reliability and validity. These latter concepts can thus be viewed as providing the linkage among, and the justification for, various practices. However, how one defines, say, the term "valid", depends on the logic of justification one adopts -- in particular, the notion of "truth" one accepts (truth as coherence or "consensus" among investigators and subjects, or truth as correspondence with some independent, objective reality) -- and not on the specific investigative techniques one employs.

*Parallel accounts of criteria for judging qualitative and quantitative research.* Another factor which has tended to blur the distinctions among the paradigms has to do with the way in which recent discussions of logic of justification issues have been framed with regard to qualitative research. Guba (1981), for example, has elaborated criteria and procedures for assessing naturalistic inquiry by trying to deliberately mirror those which are associated with inquiry. Thus, while the assumptions of naturalistic inquiry are part of an idealist epistemology, the mechanisms for distinguishing good from bad naturalistic inquiry presented by Guba require the assumptions of realism which underlie traditional science.

Let me elaborate on this last point by way of an example, following Smith and Heshius. For the quantitative-objectivist inquirer, truth value is a question of internal validity. The criterion of internal validity is a correspondence or "isomorphism" between the data or the inquirer's statements and an external, independently existing reality. For naturalism, truth value is a function of "credibility" and this is defined by Guba as the correspondence between the investigator's statements and the perceptions and interpretations, the constructed realities, of participants in the social context under study. Thus, it appears as if credibility and validity are quite similar or analogous.

But there is a problem with this elaboration of the meaning of credibility which undermines the illusion of parallelism. The use of correspondence or isomorphism as a criterion presumes that what is known -- whether that be an independent, objective reality or a constructed, interpreted, subjective reality -- has a status independent of the inquirer, and can be characterized without qualification or conditioning by the inquirer's values, interests and purposes. It is clear, then, that correspondence makes sense as a criterion only if one first accepts the epistemological assumptions of realism concerning the separation of mind and world, and the related notions of neutrality and objectivity. But these assumptions are explicitly rejected by naturalists. According to the tenets of naturalism, as we saw in Chapter 2, the inquirer can only offer an interpretation (coloured by his own purposes, values, interests) of the interpretations of others (which are similarly shaped by their own values, interests and

purposes): a construction of a construction.

Similar problems arise with respect to Guba's attempts to provide analogues for other criteria for judging inquiry, such as e.g., reliability. It seems that such ideas are simply not amenable to reconstruction within the qualitative framework. Smith (1985) provides a more concrete demonstration of this in his discussion of the problematics of the notion of test validity within the context of naturalistic assumptions.

Thus, Guba's attempts to match criteria for evaluating qualitative research with those connected with objectivist research suffer from internal inconsistencies and serve only to obscure the fundamental tensions and mutually antagonistic elements which separate the paradigms. Guba is not alone in creating this false impression, however. Lecompte and Goetz (1982) also present discussions of validity and reliability that suggest they are interpretable in the same way for both quantitative and qualitative inquiry, while Miles and Huberman (1984a) avow that their goal is to make qualitative inquiry "scientific in the positivist sense of the word" (p. 21). Paradoxically, so far from increasing the respectability and status of qualitative research, such goals can only lead to naturalistic inquiry being subverted and preempted by the assumptive framework of quantitative science.

*Compatibilism at various levels.* The idea that paradigms are not truly incompatible can also gain plausibility from focusing on comparisons at levels other than the basic assumptive frameworks or logics of justification. The potential compatibility of research traditions can be approached in four different areas: data, procedures, knowledge claims, and logic of justification or assumptive frameworks. For convenience, let us label these levels (1) through (4).

We have already allowed that paradigms may be compatible at level (2), but have also seen that this is relatively trivial and does not imply compatibility at level (4).

Quantitative and qualitative approaches may also be compatible at level (1). Howe (1988) makes much of this in his efforts to establish the thesis of paradigm compatibility. He distinguishes two senses of the terms "qualitative" and "quantitative" as they apply to data: an ontological sense and a measurement sense. In the ontological sense, data are qualitative if

they are "intentionalist", i.e., if they incorporate values, intentions or beliefs; otherwise, they are quantitative. In this sense, it certainly does not appear that the distinction marks any important boundary between naturalistic and scientific inquiry. It is a truism that scientific inquiry often does deal with data that are qualitative in Howe's ontological sense.

In the measurement sense, data are said to be qualitative if they fit a categorical measurement scheme, quantitative if they are appropriate to an ordinal, interval or ratio scale. Again, it is true that it is difficult to understand how this distinguishes alternative research traditions. Scientific inquiry does deal, sometimes, with categorical data, although admittedly there is generally a preference for data that can be submitted to the more subtle analyses offered by parametric statistical machinery. The distinction is also a somewhat fluid one. What is qualitative today, in the measurement sense, may become quantitative tomorrow, with the development of a new instrument. I am quite prepared to allow, then, that it is not any hard, simple distinction concerning the nature of data which separates the paradigms. But, again, compatibilism at this level (1) does not entail compatibility at level (4), or even at (3) for that matter. The crux of the distinction between qualitative and quantitative approaches lies with the explanations, the hypotheses, and their justifications, and not the data themselves. Quantitative researchers may collect data regarding beliefs, intentions and values, but their use of this data, and the justification of their conclusions, will not be interpretivist in the sense of naturalistic inquiry. It is misleading, then, to use comparisons at the level of data to suggest a real consolidation of the paradigms.

Turning our attention to level (3), it is perhaps conceivable that certain knowledge claims could be made, from each of the two paradigms, that would be expressed in such terms as would superficially appear to allow them to be compared and characterized as mutually consistent or inconsistent. But, still, the justification of those knowledge claims would be different. Thus, their mutual consistency or inconsistency would not carry any real evidential weight in assessing the claims advanced from within either paradigm. Alternatively, if the thesis of relativism based on incommensurability is true, then one cannot compare claims issuing from

different paradigms, even if on the surface they might appear to be stated in the same language. Either way, real or apparent, similarities between the knowledge claims advanced from within the different paradigms cannot be taken to obviate paradigm differences at level (4), or to render them insignificant.

*The case for combining methodologies.* Given the preceding comments and conclusions, it seems that there is no sense in combining methodologies (alternative level (4) specifications of paradigms) in inquiry. Once compatibilism at level (4) is rejected, then the idea of any kind of triangulation or convergence on the truth from the direction of alternative modes of inquiry is highly problematic. Suppose that the qualitative efforts support an hypothesis,  $S_1$ . And suppose further that the quantitative approach to investigating the same or closely related issues underwrites an hypothesis,  $S_2$ . Even if we allow that the two are stated in a common language, we can hardly regard them as standing in a state of mutual support (or contradiction, as the case may appear to be). From the perspective of the qualitative or interpretivist approach, the statement  $S_2$  has no warrant. And vice versa: from the scientific viewpoint,  $S_1$  is unfounded. Note that the problem as I have stated it here is not one of incommensurability. I have allowed that  $S_1$  and  $S_2$  may be stated in the same language, and so may stand in the logical relationship of consistency or inconsistency with one another. The difficulty is rather that our confidence in either  $S_1$  or  $S_2$  cannot be affected by this relationship.

The only alternative to such a conclusion is to posit some extra-paradigmatic criteria of trustworthiness that would assign equal weight to both paradigms. If we disregard the possibility of orthogonalism, for a moment, and suppose that different paradigms may address roughly the same questions, then any such criteria which were weak enough to balance the contradictory assumptions of the naturalistic and objectivist camps might have to be weak enough to lead us beyond compatibilism to terminate in an extreme form of irrationalism and relativism. (If orthogonalism is basically true, it is possible that we may then allow for extra-paradigmatic criteria that legitimate equally the different paradigms without assuming the stance of extreme relativism.)



### *Orthogonalism*

There is no question that certain modes of inquiry are orthogonal in the sense that they address different types of questions. Historical inquiry has a different subject matter from empirical science and correspondingly different methods of inquiry, different canons of evidence and inference and so forth. In other cases paradigms are just as obviously non-orthogonal. A good example here would be General Systems Theory (GST) and traditional science. GST is advanced as a replacement for the classical scientific approach, over which it claims certain advantages such as the ability to deal with intractable complexity and a greater potential for unifying the sciences (von Bertalanffy, 1968).

The question of the relation between qualitative and scientific approaches in the social-behavioural arena is not so easily answered, however. In one sense, the goals of ethnographic research as propounded by purists are largely orthogonal to empirical science. The former is idiographic, the latter seeks the broadest generalizations. Of course, the results of a qualitative study may be generalizable in certain respects. Generalizability, after all, is a question of validity and sampling. If one happens to sample one case which is very representative of a class of situations, and manages to describe that case accurately in certain respects, then one's conclusions will likely be generalizable to that class. But the methodology of ethnography is not explicitly designed to ensure generalizability.

The difficulty, though, is that qualitative research is more loosely defined in the field of education than it is in the field of anthropology, a situation which some purists find alarming (Rist, 1980; Wolcott, 1980; Fetterman, 1982). The question whether, in education and educational technology, the qualitative approach is orthogonal to the scientific one is an empirical question, to be answered by looking at how qualitative methods are actually being used, by examining the questions to which it is actually being applied.

There does in fact seem to be a trend towards viewing qualitative research as an alternative or replacement for scientific research that renders them non-orthogonal. The evidence of this is to be found, in part, in talk concerning "theory" and theory development in

qualitative inquiry. Recall that in Chapter 2 I recounted Guba's claims that theory is "more powerful" when it is derived with qualitative methods. The term theory carries strong connotations of principles of general knowledge.

But a better example, a more concrete one, can be drawn from the very recent core literature of educational technology. Neuman's (1989) article entitled "Naturalistic Inquiry and Computer-Based Instruction: Rationale, Procedures and Potential" clearly places the qualitative paradigm in the role of formulating reliable generalizations. (The author uses the term "naturalistic" as a synonym for "qualitative", in accordance with practice in the field of education and contrary to the convention I adopted earlier.) Neuman notes that in the past qualitative investigations of Computer-based instruction (CBI) have focused on patterns of implementation and social effects, as opposed to instructional issues. However, she proposes that qualitative inquiry is ideally suited to resolving as yet unanswered questions regarding the basis of instructional effectiveness in CBI. She stresses that:

courseware designers cannot yet avail themselves of a fully developed set of empirically based principles that specifically address the minute details inherent in designing interactive instructional materials. They have access to no verified taxonomy of strategies that enhance the possibility that *this* component of *this* lesson will be effective for *this* learner in *this* situation. (p. 40)

This is indisputably a call for the development of a framework of generalizations. Apparently, Neuman is misled by her understanding of the term "context-bound". What she is advocating is adherence to a paradigm which will allow one to map optimal instructional strategies against certain learner and subject matter variables. The goal, then, is an explanatory or at least predictive framework that has general validity. But she believes that the qualitative paradigm is best suited to achieving this goal because it addresses realities which are context-bound. She has apparently confused the notion of being context-bound or unique, in the sense in which ethnography builds idiographic knowledge, with the idea of a general theory or set of laws which may be operated with different initial conditions (representing different "contexts") and which may

reflect different interactions among variables of interest. In Chapter 2 I argued that this same confusion underlies the argument against the possibility of a social science which is based on problems alleged to be attached to the notion of generalization.

Stated more explicitly, her argument runs thus:

Premise 1: Qualitative methodology best deals with reality that is context-bound.

Premise 2: Any adequate explanation of the effectiveness of CBI will reflect the circumstance that the phenomenon is context-bound.

Conclusion: The best methodology for investigating and understanding instructional effectiveness of CBI is the qualitative mode.

The fatal equivocation is on the term "context-bound". In premise one, it must be interpreted as designating a situation or phenomena which is unique. In premise two it means there are many variables of interest and possibly many interactions among these variables, so that perhaps no two situations will appear identical. At a deeper level this equivocation thus rests on a confusion of description and explanation. One of the features of scientific theories is that they can explain, in a single framework, phenomena or situations which superficially appear diverse (i.e., at the more gross levels they have different descriptions).

Neuman's conflation of these two notions is apparent throughout her paper. It is perhaps most strikingly evident in a section entitled "The Nature of Truth Statements," where she provides a standard qualitativist account of truth. Reality is said to be context-bound and multiple. Quoting Guba and Lincoln (1982), she maintains that generalizations or reliable context-free statements are unobtainable. She writes: "The 'truth' of the medium is revealed not in global statements of effectiveness, but in insights into the individual experiences of particular learners interacting with particular courseware in particular settings" (p. 42). However, she then goes on to assert that: "The development both of an understanding of the interrelationships between courseware and context and of a wide array of working hypotheses about how courseware can best be used in a variety of contexts is crucial to the design of courseware that will enhance those individual experiences" (p. 42). Again, she is proposing the search for an

explanatory framework comprised of generalizations, regardless of the use of stock phrases from qualitative inquiry such as "working hypotheses" in place of "generalization".

Further misunderstandings concerning the paradigmatic assumptions of qualitative inquiry are suggested in other passages. For example, Neuman argues that the effectiveness of sophisticated branching that provides multiple learning paths in CBI reflects the idea that there are multiple realities and that these are individually constructed (p. 41). The qualitative paradigms's principle of the multiplicity of realities is generally intended to convey the notion that a person's view of the world is coloured or conditioned by the unique associations that evolve among each individual's concepts, by the connotations she attaches to them, and by her values. The fact that individuals may have different learning styles or preferences and may exhibit different background knowledge, which will influence which learning sequences are preferred or optimal (thus raising the need for effective, sophisticated branching in CBI), does nothing to support the contention that there are multiple realities and that these are essentially irreducible. In other words, the circumstances that dictate the need for branching do nothing to establish the assumptions of an idealist epistemology with solipsistic tendencies over some form of realism.

So there is evidence that qualitative inquiry in educational technology is being treated, inappropriately, as an alternative to quantitative inquiry. This is not to say that qualitative and scientific inquiry should not be regarded as orthogonal but only that, as an empirical fact, there is a tendency towards treating them as non-orthogonal. Not only has Neuman suggested that the same issues can be addressed, but also implicitly that the same kinds of answers can be sought, in qualitative as in scientific inquiry.

The basic difference between the two approaches, so far as the issue of orthogonality is concerned, lies in the direction of the answers pursued. Science seeks to uncover the significant generalizations, while qualitative inquiry is designed to capture, describe, reconstitute or exhibit what is unique. One cannot say *a priori* what dimensions of social or behavioral phenomena will admit of underlying structure that will enable the formulation of reliable generalizations. So one should expect that there will be considerable overlap in the issues

addressed by proponents of the two approaches at any point in time. In fact only the progress of science involving the development of successful explanatory frameworks will allow us to demarcate more clearly the issues that are legitimately the respective domains of the two approaches. Even then, there can be some argument that one is entitled to utilize the qualitative approach wherever one wishes essentially to describe or exhibit some truth, rather than to explain it.

In short it seems that orthogonalism is hardly an all or nothing issue so far as the qualitative and scientific paradigms are concerned. There will inevitably be some overlap among phenomena addressed, but where the goal of inquiry is a general explanatory and predictive framework then the appropriate methodology must be the scientific approach.

#### *The Arguments for Relativism*

Garrison (1986) has provided a recent summary of the basic principles issuing from postpositivistic philosophy of science which, collectively, provide the basis for relativism. He enumerates these as:

- (1) -- the problem of confirmation
- (2) -- the underdetermination of theory by logic
- (3) -- the underdetermination of theory by experience
- (4) -- the Quine-Duhem thesis
- (5) -- the theory-ladenness of experience or observation
- (6) -- the incommensurability of theories

Let me briefly examine the meaning of these, and consider their respective contributions to a relativistic stance *vis-a-vis* alternative paradigms. Afterwards, I will proceed to examine the arguments for those principles among the six identified which are most influential in this regard. Our discussion will begin with Garrison's characterization of some of these principles, but will quickly move on to question the relevance of (1) and (2) to relativism, and to provide our own detailed analyses and evaluations of principles (3) through (6), and the relations among them. Finally, I will consider and reject Garrison's solution to the threat of relativism posed by these

principles. Garrison's position seems to be that these principles are essentially correct, but that relativism can be circumvented by an appeal to common sense. On closer inspection it will be obvious that this solution is simply incoherent. Our own view is that the various arguments for (3)-(6) are insufficient to establish these principles, thereby dissolving any foundation for relativism.

*The problem of confirmation.* Recall that the standard account of theory testing, the so-called "hypothetico-deductive" model, works as follows. With the aid of certain auxiliary definitions (called rules of correspondence, operational definitions, or coordinative definitions), and certain auxiliary assumptions concerning the initial state or conditions of the system under investigation, a prediction of some specific observable event is generated, deductively, from the theory concerned. Appropriate observations and measurements are then carried out to confirm or disconfirm the prediction.

A disconfirmation is quite telling, of course. Because the basic scheme is deductive in character, a false conclusion (barring the possibility of contamination) is more or less conclusive. It ensures that some element in the theory concerned, and/or in the associated auxiliary assumptions employed in deriving testable consequences, is also false.

The results of confirmation, however, are less decisive. Again, because of the character of deductive inference, no amount of successful prediction can ever establish, categorically, the truth of a theory. In a deductive scheme of reasoning the truth of the conclusions (in the context of the hypothetico-deductive model of theory validation, the predictions) does not logically entail the truth of the premises (here, the conjunction of the statements of the theory).

What conclusion can we draw from this? Simply this: acceptance of a theory must always be tentative. But does this set one on the path to relativism? No, the fact that there is always the logical possibility that even the most well-confirmed theory will turn out to be false does not preclude the development of rational criteria to assess the relative merits of competing theories. It is not the case that we must regard every theory or hypothesis which has been confirmed, but not yet disconfirmed, as on equal footing.

For example, we may prefer the stronger of two theories (the one which asserts more, empirically), the one which is simpler in some sense, or perhaps the one which seems to afford us a faster rate of solving problems in a certain domain. We may continue to prefer the theory with such characteristics until such time as it is seriously refuted, at which point the alternatives must be reassessed.

This brings up an important point: the case for relativism derives its strengths not so much from the logical character of confirmation and that old saw, the problem of induction, as it does from attacks on the very possibility of refuting theories. These attacks arise in conjunction with the latter principles taken up below, especially various expressions of the Quine-Duhem thesis and conventionalism.

Thus, it must be concluded that the logical difficulties associated with the idea of verifying a theory do not constitute grounds for relativism.

*The underdetermination of theory by logic.* Garrison notes that, because theories are a complex form of discourse, it is frequently the case that no single set of experiments can refute an entire theory, but rather can only affect one statement of the theory or some subset of interrelated statements. It is difficult to see how this supports relativism. In the first place, it may well be that an experiment (particularly a crucial experiment) will reflect on the central portions of a theory. A good example of this might be the null result obtained by Michaelson and Morely in their experiment to test the ether hypothesis (1887). In such a case, the theory concerned is pretty well refuted. If, on the other hand, there is good reason to suppose that it is peripheral statements (auxiliary theory required for testing, statements of initial conditions, correspondence rules) that are jeopardized, then this, too, has a definite bearing on whether to reject the theory. In this case, it mitigates in favour of retaining the theory.

It is only if one supposes that one always has an arbitrary, free choice over which part of the overall theoretical structure to blame for experimental failure, the theory itself or the auxiliary assumptions which must be engaged in order to test the theory, that underdetermination of the theory by logic becomes a basis for a relativistic outlook -- for then one can always choose to

retain any particular theory and shift the blame for an apparent failure in prediction or explanation onto the auxiliary assumptions. At this point, the principle of the underdetermination of theory by logic becomes an expression of the Duhem thesis, which I will evaluate shortly.

*The underdetermination of theory by experience.* Garrison expresses this principle thus: "given any finite body of data, an infinite number of theories may be tailored to fit the body" (p. 14). He gives the familiar "best-fit" example of a plot of data, relating some independent variable,  $x$ , and some dependent variable,  $y$ . The data fall in a straight line. And it is therefore natural to assume that the choice of a single statement theory to account for these data, in the form of some mathematical function, must be the linear equation:  $y = mx + b$ . Yet there exist an indeterminate number of alternative, polynomial, equations of various order, which could also pass directly through every point on the plot. How do we justify the preference for the linear function? The spectre of relativism is raised once again.

Garrison offers two sets of considerations to deflect relativism. In the first place, he acknowledges that "simplicity" is a poor guide, here, insofar as it is somewhat relative. (What if, he asks, the data had been plotted in polar coordinates?). He fails to notice, though, that considerations of simplicity would still prove inadequate to dispel the problem of underdetermination, even if an unambiguous formulation of the meaning of simplicity could be forwarded. Unless a criterial relation between "simplicity" (defined in some unambiguous way) and "truth" could be established, the basis for preferring, all other things being equal, the simpler of two theories or hypotheses would have to be regarded as merely an arbitrary aesthetic. Despite concerted efforts, such a relationship has never been established, and the attempt has been given up as a bad cause (Barker, 1957; Bunge, 1963; Rudner, 1961; Friedman, 1972b).

Garrison believes, however, that the conclusion of underdetermination and its consequences can be escaped on the basis that data are generally interpreted within the framework of some pre-established theory, which provides guiding assumptions which can generally settle such matters as the "best-fit" dispute outlined above, nonarbitrarily. The difficulty is that the possibility of underdetermination can be extended to encompass the background



theories as well, in their entirety – a possibility of which Garrison seems unaware. The thesis of the underdetermination of theory in general is called conventionalism, universal conventionalism or global conventionalism, and it is this thesis, rather than the simple example Garrison focuses on, which poses the significant threat of relativism from the direction of underdetermination of theory by evidence or experience.

It will be convenient, at this point, to change our tack somewhat. Of the four latter principles noted by Garrison, three – the Quine-Duhem thesis (sometimes called, simply, the Duhem thesis), incommensurability, and the underdetermination of theory by experience (recast by us as global conventionalism) – have actually all been represented in the literature of philosophy as related. The focal point is global conventionalism (henceforth abbreviated as GC). The Duhem-thesis has been confused with GC and, even when recognized as distinct from GC, has also been represented as affording support for the latter (cf. Losee, 1980, p. 165). Finally, the incommensurability thesis has been explicitly referred to as "conventionalism" (Tolafson, 1982).

It will be useful, then, to approach these last principles with a view to assessing any aspect of mutual implication or support among them. It will be seen that they are actually quite independent of one another, and that the arguments for each are inadequate. Garrison's fifth principle, the theory-ladenness of experience or observation, figures prominently in the arguments for incommensurability and will be addressed in that context.

*The Duhem Thesis, Incommensurability  
and Global Conventionalism*

*Initial Conceptualization of GC*

To begin with, we need a working definition of GC, somewhat more precise than the one provided above where relativism was first introduced. GC, it has already been remarked, asserts that scientific theories are of an essentially arbitrary character. The doctrine maintains that theories as a whole or as systematic entities are merely constructs and that as such they are arbitrary in an epistemologically significant sense. The notion that theories are constructs in a

radical sense predates the modern formulation of global conventionalism. Presumably it climaxed first with Kant's analysis of the categorical framework presupposed by Newtonian physics. But Kant, of course, expressly denied that this framework of pivotal concepts which he had identified was arbitrary. Indeed, the whole point of his attempted "transcendental deduction" of the categories was to establish this framework as the only possible one, as uniquely necessary to the scientific enterprise. Conventionalism, in contrast, carries things a step further than Kant's brand of *a priorism* by insisting that theories be construed as *arbitrary* constructs.

This point of divergence cannot be overemphasized. It entails that on the conventionalist view one must allow in principle for the existence of alternative theories, theories which will be such that there can be no cognitive or methodological rationale for choosing any particular alternative over its competitors. Otherwise there is no sense to the notion that theories are arbitrary conventions. The way to broach the issue of conventionalism is patently *via* the question of the existence, or the possible existence (if none are in evidence), of such alternative theories. This much seems obvious, but nonetheless it is a matter which has not been addressed in a fully responsible manner in the literature devoted to the subject of conventionalism -- as I shall later show.

In brief, the alternative theories alluded to must satisfy two principal requirements. First, they must yield exactly the same observational consequences or, in the popular jargon, must "save the same phenomena". So the only grounds for indicating a preference for one particular alternative would necessarily have to be noncognitive ones. Such considerations might be aesthetic or pragmatic. Some alternative theory might be rather more elegant than its available counterparts, or it might contain a more efficient algorithm. But the point is that no such advantage would allow us to infer the cognitive superiority of a given theory from among a set of conventional alternatives -- at least certainly not directly. Unless this first requirement is satisfied by alternative theories one can make no sense of the notion that they are conventional, that their status is, in an essential and important sense, arbitrary. It should be remarked that this condition must be satisfied rigorously. Approximate or even virtual empirical equivalence will not suffice to

sustain conventionalism.

The second requirement which must be met by these theories is that they must be genuinely distinct. There must be no translation procedure, no mechanical operation, available for transforming a theory into one of its putative alternatives. If such an apparatus exists, then the alleged alternative theories are not alternatives in any epistemologically significant sense, and their existence does not substantiate the claims advanced by universal conventionalism. They are, in such case, not truly distinct theories at all, but rather merely trivial semantic or syntactic variants of a single theory.

Trivial semantic conventionalism (TSC) rests upon the circumstance that we can assign linguistic signs or symbols to our concepts in whatever way we choose, by an act that involves an element of definition, stipulation or convention. All forms of discourse, not merely scientific theories, are conventional in this rather uninteresting sense that their linguistic mode or form of expression might have been different from what it is. The sentences of two theories which are trivial semantic alternatives of the same theory may be different. And, if the same sentence appears in both theories, it may turn out to have different truth values. But it does not follow that they are truly distinct or arbitrary in any important sense. The old philosophical distinction between sentence (a linguistic utterance) and proposition (the meaning of a sentence) is useful here: in this situation the sentences of the theory have changed, but it still comprises the same propositions.

Much more needs to be said about these requirements. The central notions of "distinctness" and of "observational equivalence" need to be discussed more thoroughly as they are somewhat problematic. However, I prefer to defer further consideration of these topics. For the moment, I shall proceed to examine the various interpretations of the Duhem-thesis and incommensurability, and to summarize their relations to GC.

### *The Duhem Thesis*

This thesis, associated with the French physicist Pierre Duhem, is to the effect that isolated hypotheses cannot be refuted, that in science the results of observation inevitably reflect

not on any single hypothesis, but rather only on the theoretical ensemble concerned as a whole.

In *The Aim and Structure of Physical Theory* Duhem wrote:

The Physicist can never subject an isolated hypothesis to experimental test, but only a whole group of hypotheses; when the experiment is in disagreement with his prediction, what he learns is that at least one of the hypotheses constituting this group is unacceptable and ought to be modified; but the experiment does not designate which should be changed. (1974, p. 187)

Stated so baldly, the thesis seems false. The suggestion appears to be that when a theory is refuted by experimental results the theory in its totality must be blamed; it is not possible to isolate and identify some precise hypothesis or set of hypotheses to which blame may be attached from within the group of hypotheses comprising the theory. This thesis would be true in general only if the postulates of theories tended invariably to betray an unlikely and, indeed, undesirable trait, namely logical interdependence. In point of fact, the postulates of physical theories tend by design more frequently to be independent. It follows we may expect that in some cases, at least, it will be possible to single out the hypothesis, or subset of hypotheses, responsible for the failure of a theory. One can accomplish this quite rigorously by providing an independence proof, the simplest form of which consists in presenting a model which satisfies all the axioms of the theory concerned except the tainted one(s).

There is, however, another somewhat more generous interpretation which can be attached to Duhem's writings. According to this account, the Duhem thesis exploits a certain inductive latitude involved in the testing of hypotheses which results from the necessary utilization of certain auxiliary assumptions that include bridge principles or correspondence rules, statements of initial conditions, and sundry collateral theory (Grunbaum, 1960). On this interpretation, isolated hypotheses are immune from conclusive falsification not because they cannot be distinguished sufficiently from the other hypotheses contained in the theories in which they appear, but rather because their testing necessitates bringing to bear additional assumptions and further, extraneous, theoretical principles. The claim here is that fault may be

laid at the doorstep of this ancillary apparatus, rather than pinned to a particular constituent hypothesis (or, perhaps, even to the theory as a whole).

On this reading the Duhem thesis amounts to the claim that, given a theory T with a constituent hypothesis H, wherever H is threatened by virtue of the role it plays in the derivation of an observation sentence, O, which is contradicted, it (H) may be preserved by altering the set of auxiliary assumptions, A, which figure in the derivation and which are external to T.

Schematically we have:

$$\{[(H \& A) \rightarrow O] \& O'\} \rightarrow [(\exists A')(H \& A') \rightarrow O']$$

where again H stands for any constituent hypothesis in a theory, A for the set of auxiliary assumptions, O for some observation sentence which is a consequence of H and A, and O' for the true observation sentence which is incompatible with O.

This thesis appears false, also, or at the very least unwarranted. To assert that A' must exist in all cases where it is required is simply to make that claim. It is an assertion and not an argument. Yet there is no assurance on general logical grounds that the necessary set, A', will always be forthcoming (Grunbaum, 1960, p. 77). That A' will always exist is apparently something which can only be advanced as an article of dogmatic faith. Philosophies and philosophical doctrines which are thoroughly whimsical, based on presumption rather than argument, have a way of falling out of favour. So, apparently, it has been with this version of the Duhem thesis.

An alternative reading of Duhem's intentions assimilates his viewpoint to that espoused by Quine and encapsulated in the latter's assertion that "any statement can be held come what may" (Quine, 1980, p. 42).<sup>11</sup> Wedeking (1969) argues convincingly for the position that on this account the Duhem thesis reduces to the trivial proposition that there will exist some language in which H and O' are consistent. This account stops short of the assertion that O' will be nontrivially deducible in this language. There is, he points out, no reason to hold Quine to the further condition that O' must be derivable from any of the true sentences in this language (call it S') other than itself. In other words, Quine is not constrained, in defending his viewpoint, to

assert of any particular sentence in  $S'$  that it is dependent on some of the other truths of that language. This is presumably entirely as it should be, for here again there is no guarantee on general logical grounds that there will exist any consistent language in which a given observation statement can be non-trivially deduced. Now of course the objection may be voiced that no scientific theory ought to contain any observation statement which cannot be deduced from the theoretical statements of the system. I am not inclined to debate this point, however, as Wedeking replies, this is a requirement of science *per se*, and the objection simply does not speak to the matter of the language  $S'$  and the question of the strength of the requirements by which we may characterize it.

Yet another commentator in the Duhem controversy, Laudén (1965), insists that it is incorrect to assimilate the views of Quine and Duhem. He also rejects the preceding interpretation of the Duhem thesis to the effect that:  $(\exists A')[(H \ \& \ A') \rightarrow O]$ . On his view, Duhem was actually asserting a weaker thesis: "Duhem is not asserting that every hypothesis can be saved, but only that unless one has proven that it cannot be saved, then it is not falsified" (Laudén, 1965, p. 297). So according to Laudén, the Duhem thesis asserts only the following. Unless we can prove that  $\sim(\exists A')[(H \ \& \ A') \rightarrow \sim O]$ , then  $\sim O$  does not constitute a conclusive refutation of  $H$ , even if  $(H \ \& \ A) \rightarrow O$ . This is indeed a weaker thesis, and consequently not a very interesting one. In fact, it is virtually a truism.

On this interpretation, the Duhemian notion that isolated hypotheses cannot be refuted is certainly saved, for in general it will not be possible to establish that  $A'$  does not exist. Actually, it will be possible to prove this only in the case where  $H$  is flatly contradictory, or where very strong extra-logical requirements are imposed on  $A'$ . Doubtless these extra-logical requirements would have to be so strong as to constitute a way of surreptitiously smuggling in the premise that  $A'$  does not exist.

But, in any event, the thesis as it has just been framed is open to the charge of irrelevance. While it is a truism that a hypothesis cannot be conclusively refuted unless it can be proved that there is no way to save it, the scientist is concerned with something somewhat more

pedestrian than "conclusive" refutation. In general, he is satisfied with a refutation which is context dependent, the context being constituted by the theory in which the refuted hypothesis occurs. He is, I think, entitled to be satisfied with such a form of refutation, especially as there seems to be a complete lack of support for any of the stronger versions of the Duhem thesis (Indeed, as a practical matter he would be justified in accepting this form of refutation even if the strongest form of the Duhem thesis, which asserts that any hypothesis can be saved, could be established. In that case the notion of "conclusive" refutation would doubtless have to be abandoned as unintelligible, but the point is that the scientist is concerned with advancing and improving his science and not, first of all, with "saving hypotheses" at any cost.)

It appears, then, that the Duhem thesis holds little interest. The various possible interpretations of the thesis which I have examined have revealed themselves to be false, unwarranted or trivial. None of this, however, has any genuine implications for global conventionalism. The Duhem thesis, in its non-trivial form, asserts something to the effect that any hypothesis can always be saved. It asserts that compensatory adjustments (adjustments which will enable us to retain the hypothesis concerned) can always be made within the body of auxiliary assumptions and collateral theory which are required in order to generate testable consequences from a theory. One might think, at first, that this amounts to an assertion of global conventionalism. But it does not. There is no version of the Duhem thesis which of itself entails the view that distinct theories may exist which are empirically indistinguishable.

To clarify this point, let us consider an imaginary situation. Suppose we have a theory T which contains a finite number of postulates. And suppose that one of these hypotheses, in conjunction with the necessary auxiliary assumptions and collateral theory, yields a prediction which is refuted by subsequent observations. If some change in the auxiliary apparatus can "save" that hypothesis, then the Duhem thesis will have been vindicated in this particular instance. But it by no means follows that there exist distinct theories which handle exactly the same data. This would not be assured even if it were possible that the theory could also be salvaged by replacing the problematic hypothesis with another. In that case we would have the

original theory T with a new set of auxiliary hypotheses A' alongside a new theory T' (T, but with H replaced by H', combined with the original set, A, of auxiliary hypotheses), the two being, let us generously allow, empirically indistinguishable.

On the face of it, these are two distinct theories. But (T & A') need not necessarily be distinct from (T' & A). They might only be semantic alternatives, especially in view of the fact that the body of auxiliary assumptions in question includes the rules of correspondence or coordinative definitions which fix the sense of at least some of the terms of the theories. In this case it is not even certain that T and T' need be said to be distinct, since we may decide to include the rules of correspondence as an integral part of the theories, rather than as external assumptions. This is a reasonable measure: it is somewhat arbitrary to insist on excluding semantic rules from the theory for the purposes of philosophical analysis. In fact, this is an understatement. Scientific theories, after all, are not uninterpreted calculi. There is just no such thing as a scientific theory without semantic rules. All this points to another problem with the Duhem thesis in its supposedly nontrivial guise, for one can hardly say that an hypothesis has been "saved" if its sense has been altered.

In brief, then, no argument for the Duhem thesis, in any of its forms, necessarily works for global conventionalism. Likewise no criticism of the thesis is bound to impugn universal conventionalism. Neither does global conventionalism imply the Duhem thesis, for the former asserts that there are distinct theories which handle the same data and this does not entail that any arbitrary hypothesis can necessarily be saved. Evidently the two doctrines are quite distinct. And, finally, it does not appear that the Duhem-thesis can be established in a way which entails any significant (radical) form of relativism.

*The Weltanschauungen View: The Incommensurability of Paradigms*

This is the view of science sometimes labelled conventionalism but more aptly referred to as the "Weltanschauungen" view. The chief purveyors of this view have been T. S. Kuhn (1970) and P. Feyerabend (1975). At the center of the weltanschauungen perspective is the incommensurability thesis, which we have encountered, briefly, above. As I have already



remarked elsewhere, under its usual interpretation the weltanschauungen view is entirely at odds with the basic precepts of GC as represented in this dissertation, because of its dependence on incommensurability. However, it is possible to formulate a modified version of GC within the constraints set by the basic conceptual limits of the weltanschauungen view, and I shall examine this possibility, also.

The weltanschauungen school offers a radical interpretation of the growth and development of scientific knowledge. In its Kuhnian incarnation it asserts that the evolution of scientific knowledge does not proceed primarily as a continuous and cumulative process. Rather, science is said to advance by means of a series of catastrophes constituted by "conceptual revolutions", in which old methodologies and conceptual frameworks ("paradigms") that have proven no longer viable are sloughed off by the scientific community to be superseded by radically new ones. The differences between the sciences which exemplify distinct paradigms are regarded as so extreme that they are said to be incommensurable.

The incommensurability thesis lies at the very heart of the Kuhn-Feyerabend interpretation of science. Kuhn seems to believe that a consideration of the history of science is alone sufficient to support his views (Kuhn, 1970, p. 4). However, his critics have rightly countered that the very interpretation of the history of science, of conceptual change within the discipline, must involve some prior assumptions of a philosophical nature. In this connection it is worth remarking that the conceptual revolutionist's bible, Kuhn's *The Structure of Scientific Revolutions*, contains surprisingly little in the way of carefully detailed case studies. A good deal, perhaps the greater proportion, of his argumentation proceeds from general considerations in epistemology and the philosophy of language, and anecdote.

In fact, the argument for incommensurability sits squarely on the meaning-variance thesis, which is to the effect that the meaning of the terms appearing in different theories is determined entirely by the syntactical rules and usage implicit in these theories. So according to this view the meanings of terms employed in science are entirely relative to the theories in which they appear. The meaning-variance thesis, if swallowed whole, is sufficient to ground

incommensurability, or at least it is when taken in conjunction with another central tenet of the weltanschauungen position which is to the effect that the sharp distinction which empiricists have traditionally posited between observational and theoretical terms is not sustainable. Hence the meaning-variance thesis is construed as extending to the entire vocabularies of theories. On this view there are no extra-logical terms whose meanings are preserved through the transition from one theory to another.

Couched in these terms, the meaning-variance thesis thus contradicts the idea of a neutral observation language, of a separate vocabulary which remains semantically stable despite changes in theory. Without presupposing such a predominantly neutral observation language, theories must appear as incommensurable: there can simply be no relations of consistency, incompatibility, relative confirmation or reduction (deducibility) among them.

Needless to say, radical incommensurability is completely at odds with the notion of GC as it has been previously characterized, for the latter has been said to address itself to distinct theories which are empirically equivalent. If there are no theory-independent phenomena then plainly our theories cannot be underdetermined by the phenomena: the entire notion is devoid of sense. Within the framework of the weltanschauungen view the very idea that two distinct theories are "about" the same phenomena is simply beyond the pale.

Kuhn's views are thus not intended to be taken as a form of conventionalism, at least not on his own terms. Kuhn addresses himself to the history of science. He sees it as advancing in the following fashion. At some point an accepted theory suffers a crisis. It is no longer able to save the phenomena it is "about". Subsequently there is a conceptual revolution, the scientific community adopts a new paradigm, and a new theory is developed which is incommensurable with its predecessor. The new theory, contrary to the tenets of conventionalism, can be preferred on epistemic grounds, even though there is no neutral set of facts which could be referred to in order to make a comparative assessment of the two. This is because the first theory has failed the phenomena on its own terms, while presumably its successor has not yet evidenced that defect.

So Kuhn is appealing to the history of science, claiming that each major advance in the discipline occurs as a result of this process which consists in a theory undergoing a crisis, becoming no longer viable on its own terms, and then subsequently being supplanted by a new theory embodying a novel paradigm. And to this extent, Kuhn is not saying anything that implicitly carries a commitment to conventionalism. However, the question remains whether a form of GC may be compatible with the weltanschauungen analysis of science, whether one can make sense of the notion of GC set in a context in which there is said to be no stable observation language.

The answer, it seems, is affirmative. Let us suppose that Kuhn and his advocates are right concerning the incommensurability of theories. Then imagine a situation in which a theory embodying a new paradigm is advanced by a part of the scientific community, while there is an accepted pre-existing theory in place which has not yet fallen into a crisis. By hypothesis, these theories would be incommensurable and the choice between them would have to be regarded as a conventional or nonepistemic one.

It is possible, then, to make sense of the notion of a form of GC within the constraints of the weltanschauungen view. However, I shall still proceed to regard conventionalism as the doctrine which asserts that there may exist theories which are observationally equivalent but distinct, a definition which presupposes a neutral, stable, observation language and which locates the potential for conventionalism in the theoretical portion of our theories.

The justification for this approach is simply that the incommensurability thesis is false. It has been pretty well demolished by a number of arguments. In the first place, there have been carefully detailed analysis which have shown the commensurability of such theories as those of Newton and Einstein (e.g., Friedman, 1977; Angel, 1980). These case studies tend to presuppose a framework of referential semantics which seems entirely appropriate. The arguments for incommensurability trade, it seems, in part on an equivocation between two components or types of meaning: sense and reference. The sense we attach to the terms which designate observable properties and processes can certainly be modified or informed

somewhat by the way they are subsumed under a particular conceptual (theoretical) framework; but this does not dictate that we can have no assurance that the reference of our observation terms is preserved in the shift from one theory to another.

Of course, the incommensurability thesis also draws on the claim that no useful or legitimate distinction can be drawn between theoretical and observational terms. But, as I argued earlier, the arguments against the observational-theoretical language distinction do not succeed.

It is also worth noting that the meaning-variance thesis, asserting as it does that the meanings of terms are theory-dependent, runs counter to the intuitive notion that terms have meaning within a language, not within a theory. Consequently, the thesis has highly paradoxical implications if it is taken seriously. Let us suppose, for example, that we have a theory,  $T$ , which yields a prediction,  $O$ , which turns out to be false. It appears that if we push the meaning-variance thesis to its logical conclusion, then we cannot really say that the theory has been refuted, for the sentence  $O'$  or  $\sim O$ , which we would normally say contradicts the theory, is not commensurable with the theory or its derived consequence,  $O$ .  $\sim O$  is not a sentence belonging to the theory  $T$ ; it is rather a sentence in the pretheoretic "observation" language. On the meaning-variance thesis, it is meaningless from the perspective of the theory.

Hence, if one accepts the meaning-variance thesis completely, one is pushed ultimately to the position that a theory can never be refuted, not even on its own terms, short of revealing itself to be internally inconsistent. This does serious violence to the notions of meaning and of refutation and hence constitutes an effective *reductio ad absurdum* of the position. It may be objected that if  $O$  is a consequence of  $T$ , then  $\sim O$  must be in the language of  $T$ , also. Granted,  $\sim O$  is in the language of  $T$ . But the point is this: we do not know, so long as we assume the meaning-variance thesis, that the terms occurring in  $\sim O$  have the same interpretation as they do in  $O$ . Again, this is because in the context of the *weltanschauung* view we do not properly speak of meaning within a language but rather of meaning within a theory.

In conclusion, we have seen that it is possible to formulate a version of GC which is

consistent with the incommensurability thesis, the principle of the theory-ladenness of observation, and the meaning-variance thesis which, together, constitute the core of the weltanschauung view of scientific discourse. However, these basic ideas are not sound, and so, once again, I find no basis on which to advance relativism. The incommensurability thesis, I noted, is an alternative basis for relativism to GC. However, it, too, trades on the ideas of meaning-variance and theory-ladenness which I have just rejected.

#### *GC and Criteria of Theory Synonymy*

At this point I will shift my efforts to a closer examination of GC itself. We begin by considering various possibilities for advancing a more precise statement of the meaning of GC. The key will be to provide a more exact criterion for the assertion that two observationally equivalent theories are distinct. To this point I have relied on the intuitive notion that two theories are distinct providing that they are not merely trivial semantic alternatives. This is not very edifying, however, unless a more exact account TSC can be furnished. I shall argue that two theories are trivial alternatives if and only if they are isomorphic (in a special context). To this end I shall first try to discount both weaker and stronger requirements for theory synonymy.

A word about the tools I will be using is in order before we proceed. In order to clarify the concept of GC, it will be useful to employ the formal concepts of "theory", "interpretation" and "model". In the formal sense, a theory is a linguistic entity comprising a set of statements in a well-defined language. Generally a theory will be expressed in a first-order language, such as the first-order predicate calculus, and will be cast in axiomatic form. A theory has no meaning, it is all syntax until it is fitted out with a semantics or "interpretation". An interpretation is a set-theoretic entity rather than a linguistic one. Technically, it comprises the following elements: a domain,  $D$ ; for every sentence letter in the language, a truth-value; for every  $n$ -place predicate letter in the language, a "characteristic function",  $O$ , with  $n$  argument-places; for every function symbol in the language, a function on the domain,  $D$ ; for every name in the language, a designation in the domain. A "model" is then defined as an interpretation in which the sentences of the theory are true.

The advantages of using these formal conceptions are the clarity and rigour which they may lend to our discussion. It is also possible that GC may be construed as an analytical rather than a synthetic claim: that is, as a formal claim about the relationship among a class of theories which are characterized by certain features (whatever necessary conditions can be identified of scientific discourse in general), and which share a submodel (specifically, an interpretation of the observable world) while also being non-intertranslatable or distinct (according to well-defined criteria). In that case one might hope for either a formal existence proof, establishing that such distinct but observationally equivalent theories must exist, or a formal proof of the impossibility of such a class of theories. In fact, there would even be some justification for a purely syntactical approach, as opposed to a model-theoretic one. The basic requirement is for a distinction between an observational and a theoretical vocabulary, and this demarcation can be made without recourse to the apparatus of interpretations and models. However, even in the absence of a formal proof, and even if GC is in fact a synthetic claim, the concepts introduced above can be useful in striking a more rigorous, precise account of the meaning of GC that will allow us to measure its plausibility more judiciously.

The use of model-theoretic techniques in philosophy of science is well established and includes the treatment of questions such as the semantics and logical characteristics of specific theories (e.g., Sneed, 1971; Beth, 1961) as well as the relationships among theories.<sup>12</sup> The two principal objections which come to mind regarding the use of these tools are easily met. The first is that scientific theories are not generally cast in formal axiomatic form. Clearly, however, many scientific theories are sufficiently elaborated to be rendered in axiomatic form. Several examples of this exercise can be found in the literature in areas as diverse as economics, particle physics, mathematical learning theory and biology (e.g., Suppes 1957, 1959; Woodger, 1957; Estes & Suppes, 1959; Adams, 1959). The second is that scientific theories are living, evolving cultural artifacts, and that their dynamic aspects are lost in the move to a rigid, formal axiomatic mode of representation. There is a point to this objection; however, it is not really germane to our present purposes. I am not arguing that the preferred mode of expression for

science should be a strictly formal one. I am only claiming that for the purpose of assessing the cognitive status of theories it may be useful, especially in the context of an appraisal of GC, to allow that scientific theories can be couched in formal terms that capture their structure and import at a given stage of development.

With this justification in mind, we can continue. To begin with I shall consider briefly two other requirements for theory synonymy which might appear to carry a certain amount of initial plausibility. The first is based on the notion of a match. The idea behind this criterion is that two theories are distinct providing they have different domains. The second possibility I shall explore is the notion that two theories are synonymous (are semantic alternatives) in the event they share a model.

Let us first consider a more precise formulation of conventionalism that invokes as a condition the account of distinctness which is couched in terms of the concept of a *match*. (An interpretation,  $I$ , "matches" a set,  $R$ , of quantifier-free sentences if: (a)  $I$  is a model of  $S$ , and (b) if any terms (names or function symbols) occur in  $S$ , then each object in the domain of  $I$  is the denotation of some such term. (A match is thus an interpretation in which every object in the domain is referred to -- but not necessarily named.) We have the following:

- (i)  $T_1$  and  $T_2$  are two theories.
- (ii)  $L$  is the observation language.
- (iii)  $I_L$  is the standard interpretation of  $L$ .
- (iv)  $T_1$  and  $T_2$  have exactly the same observational consequences; that is, for any sentence  $S$ , if  $I_L$  is an interpretation of  $S$  then  $(S \in T_1 \leftrightarrow S \in T_2)$ .
- (v) The languages  $L_1$  and  $L_2$ , (of  $T_1$  and  $T_2$ , respectively) each includes  $L$ .
- (vi) Terms occur in the sentences of  $T_1$  and  $T_2$  which are not interpreted by  $L$  (i.e., there are theoretical terms).
- (vii) There is no set of objects  $D$ , such that some interpretation  $I_1$  with domain  $D$  matches  $T_1$ , and some interpretation  $I_2$  with domain  $D$  matches  $T_2$ .

This list is not intended to exclude any further conditions on  $T_1$  and  $T_2$  which might reasonably

apply. In particular, it is not to be construed as preempting any conditions they might have to meet, in addition to those specified, in order to qualify themselves legitimately as examples of "science". I shall assume that there are no objections to conditions (i) through (vi), such as they are; they merely formalize the requirements of two theories which share an observation language and which make exactly the same predictions in that language. My intention here is to focus on (vii).

Condition (vii), on reflection, is unacceptable. It constitutes too strong a requirement in two different respects. First of all it necessitates that  $T_1$  and  $T_2$  exhibit the same number of terms. This is somewhat in keeping with the spirit of an idea, advanced at one time by both Putnam (1974) and Glymour (1970), to the effect that theories cannot pose conventional alternatives unless they are equivalent in such matters as descriptive simplicity. However, this idea strikes me as completely misguided. Surely all we require for conventionalism is that two theories can be observationally equivalent, but distinct. This, of course, leaves open the possibility that there might be obvious pragmatic advantages attached to one particular alternative. But such advantages, I would maintain, can never translate into epistemic superiority; they cannot be called upon to push aside the problem of underdetermination, should it genuinely arise. If this is sound, then there is no possible warrant for insisting that theories which exemplify GC contain an identical number of terms. Yet this condition is implicit in (vii) as it stands.

In addition, the requirement that two theories which instantiate the doctrine must involve different domains is itself apparently too strong, also. Presumably we would not want to call two theories which exhibited the same domain, but which described that domain differently (suppose, for example, that they contained different relations on that domain), trivial alternatives. But on requirement (vii) as stated, two such theories would fail to qualify as significant conventional alternatives. Of course, it may be that no two theories with the same observational consequences and the same domain (which also satisfied all the requirements of science *per se*) would likely be distinct in this other sense. But we do not appear to have any absolute



guarantee of this, and so there is no real justification for entering the thesis as a premise in our discussion -- however intuitive it may seem.

Now it is tempting to suppose that the first difficulty associated with (vii) might be circumvented by withdrawing the employment of the concept of a match. Suppose we retain the form of (vii) as it stands, but substitute "model" for "match". We then obtain the following weakened requirement, (vii)': "There is no domain  $D$  such that for some interpretations  $I_1, I_2$ , with domain  $D$ ,  $I_1$  is a model of  $T_1$  and  $I_2$  is a model of  $T_2$ ."

This gets around the first objection outlined above to the original formulation (vii), namely that it unjustifiably required significantly conventional, alternative, theories to exhibit the same number of terms. But it does not meet the second objection that was raised. Moreover, it generates its own peculiar problems. For instance, we can generate the required domain  $D$  trivially in every case: just let  $D = \{D(I_1) \vee D(I_2)\}$ . Thus, there is no apparent way to salvage the approach followed in (vii) and (vii)', an approach which focuses exclusively on the domains of theories which might be nominated as significant conventional alternatives.

Let us proceed, then, to consider another possible approach. Glymour (1977) has suggested in passing that two theories might be considered synonymous on the condition that they share a model.<sup>13</sup> Correspondingly, we would say that two theories are distinct only if they do not share any model. Intuitively, this criterion doesn't seem satisfactory either. It appears to be too weak. If two theories satisfy only the requirement that they share a model, then they can certainly be used to say the same things. But unless the theories concerned are both "categorical" -- and it is far from certain that our empirical theories would be, since any consistent categorical theory must have a finite domain -- there is no assurance that they can necessarily be employed to say all the same things. A theory  $T$  is said to be categorical if any two models of  $T$  are isomorphic.

Consider the case where a theory  $T_1$  is embedded in another theory  $T_2$ .  $T_1$  and  $T_2$  must then share a model. But one would hardly want to say that they must be synonymous, for  $T_2$  is richer in structure and consequently may be interpreted so as to convey information that simply

cannot be modelled by  $T_1$ . On this basis it seems that we must also reject the notion of "sharing a model" as a sufficient criterion for theory synonymy.

Now we can easily anticipate a possible objection to this dismissal. It would run something like this: "If two theories share a model, but satisfy no stronger criterion such as isomorphism (say  $T_1$  is embedded in  $T_2$  and they share a model,  $M$ ) and they are empirically equivalent, then we may expect that  $T_2$  contains some additional structure which is not being put to work, as it were. This extra structure may be described as so much excess "metaphysical baggage", and it should not be possible for such excess baggage alone to differentiate between two theories. But if we do reject the criterion of synonymy that invokes the idea of simply sharing a model and insist instead on some stronger requirement -- perhaps that they share all their models -- then this is precisely what may happen.

In fact, we could produce pairs of theories which we would have to acknowledge as legitimate instances of GC by the following trivial method. We simply take an existing theory  $T$  and append certain axioms to it, being careful to ensure that they cannot facilitate the derivation of any new observation theorems -- so that they cannot in any way affect the empirical commitments of  $T$ . This is easy enough to accomplish: one simply uses a disjoint vocabulary for the appended axioms. The original theory  $T$  and the amended version with the extraneous baggage, call the latter  $T^*$ , might then assert a legitimate claim to status as significant conventional alternatives, for we could easily construe  $T^*$  in such a way as to ensure that it would have a model which was not also a model of  $T$ ."

There is certainly something to this objection. However, I do not believe that the proper response is to allow the notion of sharing a model to serve as an account of theory synonymy. Rather, the appropriate way to deal with the problem addressed in this objection is through strengthening the requirements that a theory must meet in order to qualify as genuinely scientific discourse. We impose the requirement that our theories be suitably austere, that they carry no such purely metaphysical baggage. To this end we may insist that theories betray no theoretical principles which do not figure indispensably in the derivation of some theorem in the observation

language. This may be expressed, equivalently, as the requirement that our theories be cast in a form that is "empirically minimal". The notion of empirical minimality is owed to van Fraassen, and he defines it thus: "We may call a theory empirically minimal if it is empirically equivalent to all logically stronger theories -- that is, exactly if we cannot keep its empirical strength the same while discarding some of its models" (1980, p. 68).

There might well be some objection to this reply. Van Fraassen himself, after pointing out that "sophisticated" theories invariably reveal some extra baggage, argues that empirical minimality is not to be praised as a virtue. He says that empirically minimal theories, so far from being required, are less preferable than their more complex counterparts. He justifies his stance essentially by contending that this extra baggage which is involved has "potentialities for future use" (1980, p. 69).

But this resistance can be met quite effectively. There are two points to be considered in this regard. First, our requirement that theories be stated in an empirically minimal form is not necessarily to be interpreted as a prescription to the scientist to eschew any theory which contains extra baggage. From the standpoint of the pragmatics of theory development, such theories may offer certain advantages. However, for the purposes of conceptual analysis or philosophical appraisal, we may still insist that theories be cast in empirically minimal form. The point is not that such extra metaphysical baggage is necessarily useless, but rather that it may be methodologically misleading. Our tendency to construct theories with optional content may be more significant from the point of view of cognitive psychology and pragmatics than from a methodological perspective.

Second, the rejection of extra-theoretical baggage, at least so far as the matter of the epistemological analysis of our theories is concerned, may be defended by the following considerations. Van Fraassen grounds a preference for more complex theories in the potential for future use that may be inherent in the additional structure which they possess. The question arises, however, why should we favour these more complex theories when we have no assurance before the fact that the extra baggage they are carrying is precisely the sort of extra

baggage which will invariably prove valuable. It appears, therefore, that there is a strong case to be made for the requirement of austerity. This is certainly so with respect to the trivial case I posed earlier. The type of metaphysical baggage that  $T^*$  carries can always be introduced as it is required.

In any event, it seems preferable to exclude such pairs of theories as  $T$  and  $T^*$  from consideration as significant conventional alternatives by this strategy just adumbrated, rather than by choosing the alternative solution presented by the "sharing a model" criterion of synonymy. The extra baggage packed into  $T^*$  may well turn out to be entirely superfluous; we may never find a purpose for it. But if one is intent on declaring  $T$  and  $T^*$  synonymous then one must be prepared to label the extra baggage belonging to  $T^*$  as meaningless. This course appears to set us on a route towards some form of a verificationist account of meaning, and this is currently not a particularly attractive position. Irrelevance is not to be confused with meaninglessness.<sup>14</sup>

We require, then, a stronger criterion for theory synonymy than that afforded by the condition of sharing a model. Such a criterion is now suggested, rather naturally, by the condition that synonymous theories share all their models; that is, for any synonymous theories  $T_1$  and  $T_2$ , every model of  $T_1$  will be a model of  $T_2$ , and conversely. This condition will be satisfied just in case the two theories are isomorphic. We can then specify a routine or mechanical procedure for converting  $T_1$  and  $T_2$ , and vice versa. The procedure consists in the specification of a "dictionary": a set of sentences,  $S$ , of definitional form, which will enable us to perform the translation procedure. The procedure works as follows. Given two theories  $T_1$  and  $T_2$ , the set of sentences,  $S$ , would be appended to  $T_1$  to form a new theory,  $T_2'$ .  $T_2'$  would be an extension of  $T_1$ . From  $T_2'$  we could generate as a set of theorems all the sentences of  $T_2$ . We could then drop all sentences of  $T_2'$  containing non-logical symbols of  $L_1$  (the theoretical vocabulary of  $T_1$ ). The remaining theorems would be the sentences of  $T_2$ . An analogous procedure would of course transcribe  $T_2$  into  $T_1$ .

We can now say that a theory  $T_1$  is distinct from another theory  $T_2$  if either (or both) of

two conditions obtain: (a) On translation  $T_1$  is found to contain sentences which are incompatible with sentences occurring in  $T_2$ , or (b) on translation one of the two theories is found to contain sentences which are not expressed in the other. Note that this requirement of intertranslatability for theory synonymy does not commit us to the view that a theory may be strictly synonymous with one of its own extensions.

We need one final qualification to complete our characterization of GC. If underdetermination is just a temporary condition of an evolving theory, if we allow that one of a set of theories satisfying the criteria we have advanced so far may establish itself over its competitors by admitting of further expansion and development where they may not, then we must conclude that theories are not underdetermined in any interesting sense. One of the conditions which we must impose on two theories,  $T_1$  and  $T_2$ , which are presented for candidacy as an instantiation of GC, then, is the following: For any successful extension of  $T_1$  (call this  $T_1'$ ), there must exist an extension of  $T_2$  (call it  $T_2'$ ) such that  $T_1'$  and  $T_2'$  are observationally equivalent. And conversely.

This is obviously a very strong condition given that  $T_1$  and  $T_2$ , as well as any extensions they may spawn which are also observationally equivalent, must qualify as distinct. Certainly we would expect the condition to be satisfiable in the case where the initial theories  $T_1$  and  $T_2$  are isomorphic. In that case, they begin with exactly the same structure and hence precisely the same resources for modelling the phenomena to begin with. However, we have argued that two theories which are isomorphic are not distinct in the required sense (assuming empirical minimality and so forth, on both counts).

To summarize: I have, in this section, defended the view that isomorphism or complete intertranslatability is the sufficient condition for the synonymy of theories which are observationally equivalent and which satisfy certain other conditions which have been specified. I have now identified the following conditions which must obtain if GC is sound:

- (i)  $T_1$  and  $T_2$  are two theories which are consistent and empirically adequate.
- (ii)  $L$  is the observation language.

- (iii)  $I_L$  is the standard interpretation of  $L$ .
- (iv)  $T_1$  and  $T_2$  have exactly the same observational consequences; that is, for any sentence  $S$ , if  $I_L$  is an interpretation of  $S$  then  $(S \in T_1 \leftrightarrow S \in T_2)$ .
- (v) The languages  $L_1$  and  $L_2$  (of  $T_1$  and  $T_2$ , respectively) each includes  $L$ .
- (vi) Terms occur in the sentences of  $T_1$  and  $T_2$  which are not interpreted by  $I_L$  (i.e., there are theoretical terms).
- (vii)  $T_1$  and  $T_2$  are not isomorphic (i.e., the theories are distinct.)
- (viii)  $T_1$  and  $T_2$  are each empirically minimal.
- (ix) For any sentence  $S$ , if  $I_L$  is an interpretation of  $S$ , and  $S \in T_1$  ( $S \in T_2$ ) then there is a minimal subset of axioms of  $T_1$  (of  $T_2$ ) which is sufficient for  $S$  and which includes some sentence or sentences not interpreted by  $I_L$ .
- (x) For any successful extension of  $T_1$  which increases the empirical content of the theory (call the extended theory  $T_1'$ ) there is an extension of  $T_2$  (call it  $T_2'$ ) such that  $T_1'$  and  $T_2'$  satisfy conditions (i) through (x), and conversely.

#### *Some Conceptual Confusions*

Now that these conditions have been established, we begin to see more clearly that GC does not really possess the strong initial plausibility which seems to be presumed in much of the literature on the subject. Clearly, there are no two theories in existence which satisfy these conditions. Nor are we compelled to affirm the possibility of constructing such theories on any logical grounds. In fact, intuitively GC now looks a rather dubious proposition. Yet conventionalist literature, by and large, proceeds on the assumption that if such theories do not actually exist, they could readily be constructed. This, in the absence of genuine support, is a vast presumption and one which the anti-conventionalist is fully entitled to refuse.

Oddly, many anti-conventionalists who have written on the topic seem themselves to have been prepared to grant that GC has a note-worthy *prima facie* case. In this respect they have simply been too obliging. Their generosity can only be attributed to a failure to pay strict enough attention to the meaning and the implications of the doctrine. This is just to say that GC

has not been stated with sufficient clarity and rigour. There is considerable conceptual confusion operative in the literature which has also served to obscure the real issues.

A few additional words about these confusions are in order. For one, arguments which really support nothing more robust than trivial semantic conventionalism (TSC) are frequently mistaken as support for GC. In fact, GC is even conflated with TSC in a most serious fashion

The most striking example of this occurs in Horwich's article, "How to Choose Between Empirically Indistinguishable Theories" (1982). Horwich asks us at one point to consider two total theories,  $T_1$  and  $T_2$ , which include atomic physics.  $T_1$  and  $T_2$  differ only in that each occurrence of the term "proton" ("electron") in  $T_1$  is replaced by the term "electron" ("proton") in  $T_2$ . Thus, Horwich concerns himself with what he believes are "those cases of underdetermination which involve incompatible, empirically equivalent, *isomorphic* total theories" (p. 65). He regards this problem as a significant one and proceeds to argue that a non-conventional selection of the standard total theory  $T_1$  may be established by an appeal to our standard reference-fixing practices. He writes: "Our adoption of a whole theory formulation will constrain the referents of its terms in such a way that the alternatives will violate the requirements of our reference-fixing practice, and can therefore be rejected *a priori*" (p. 63). But, one may ask, why should one adopt one particular whole theory formulation *a priori* as opposed to any other, to begin with?

Horwich also errs seriously in maintaining that  $T_1$  and  $T_2$  are necessarily incompatible but underdetermined.  $T_2$  is incompatible with  $T_1$  only on the assumption that the reference of the terms "electron" and "proton" is held constant. However, if it is then it is a gross assumption that  $T_2$  will be empirically equivalent to  $T_1$ . In fact, if  $T_1$  is assumed to be empirically adequate, then  $T_2$  may well fail to be true to the phenomena. Horwich calls these two theories "potential notational variants" (p. 67). So long as they are potential (as opposed to actual) notational variants, they may be incompatible; but then they may not necessarily be underdetermined. And as soon as they become actual notational variants, by suitable reinterpretation in  $T_2$  of the terms concerned, the problem no longer warrants serious consideration. A clearer case of TSC we

could not ask for.

Horwich's error may also be diagnosed at a deeper level. He is apparently treating the terms "electron" and "proton" as what Kripke (1980, p. 15) calls "rigid designators". While the position of the terms within  $T_1$  has been altered in order to produce  $T_2$ , they have supposedly retained their original reference -- hence they are functioning as rigid designators. But this hardly makes sense; theoretical terms are simply not amenable to treatment as rigid designators in this context. Theoretical terms get their import from two sources: their implicit definition (that is, their relation to other theoretical terms occurring in the postulates of the theory concerned), and (in some cases) *via* their association with observational terms through correspondence rules. If one simply interchanges the position of two terms within a theory as Horwich has done, then one unavoidably alters their implicit definitions, and hence their sense. Syntax is essential in determining the import of theoretical terms.

Moreover, on the assumption that the two theoretical terms involved may occur individually in certain postulates, one may well alter the phenomena to which certain law statements are directed by interchanging them. Viewed as rigid designators, we must assume that the terms will carry with them any correspondence rules in which they appeared in the original theory formulation. Thus, as I have suggested, it is by no means self-evident that the new theory will be empirically equivalent to the original.

There are other dimensions to the conceptual confusion inherent in the literature. If Horwich's article is a prime example of TSC masquerading as GC, then on the opposite side of the coin we encounter instances where GC is dismissed as a true but entirely trivial set of claims.

To further confound matters, GC is often mistakenly identified with a defunct programme in the philosophy of science, namely, instrumentalism. Instrumentalism is the doctrine which holds theories are to be regarded simply as instruments of prediction or as "inference tickets", and which asserts that the meaning of a theory is exhausted by its observational or empirical consequences. According to this view the theoretical terms postulated by science are just useful fictions; their value is merely pragmatic or heuristic. Instrumentalism is contrasted with scientific



realism, which asserts that when a theory is well-confirmed one is obliged to affirm the existence of the theoretical processes and entities which it postulates.

The logical boundary between conventionalism and instrumentalism is not difficult to establish. Adherence to conventionalism does not force acceptance of instrumentalism for the following reasons. Suppose we are confronted with a pair of theories which are observationally equivalent but distinct. It is always possible that the ontology presented by one theory (i.e., by the theoretical portion of the theory) is the true ontology of the world, whereas the second theory is simply a false theory (to the extent that it embodies a false ontology) which happens coincidentally to be an effective instrument for generating predictions concerning observable phenomena.

Neither does acceptance of instrumentalism force a commitment to conventionalism. One can embrace the instrumentalist interpretation of the status of the status of the theoretical portion of science, to the effect that the role played by theoretical terms is purely methodological, without thereby committing oneself to the further view that there can exist distinct theories (*qua* instruments) that handle exactly the same data. In short, one can advocate instrumentalism and stop short of conventionalism.

Conventionalism and instrumentalism are therefore logically distinct. This is not to say, of course, that they are incompatible. However, it is important that they be distinguished, particularly as instrumentalism no longer enjoys much favour. To conflate the two is to obscure the real nature of the claims advanced by instrumentalism and to raise the danger that it may be rejected for the wrong reasons. That the two doctrines are often confused can be seen from an examination of philosophical dictionaries. In Angeles (1981), the entry under "conventionalism" urges us to "compare with instrumentalism", while Lacey (1976) declares that "Conventionalism is close to instrumentalism . . ."

#### *Empirical Equivalence*

A few comments concerning the notion of empirical equivalence are also in order, since we must face the possibility that GC might be incoherent on the basis that the notion of

empirical equivalence, which is crucial to the very concept of underdetermination, may be unintelligible. M. Gardner (1976) is one commentator who has explicitly used this gambit in an attempt to dispense with the problem of conventionalism. To this end he employs certain of the standard arguments against the observational-theoretical dichotomy. My position in this dissertation is clear. I have adopted the stance that the distinction stands, and that it must play a pivotal role in any acceptable formulation of GC. The standard repertoire of arguments against the distinction includes the following: arguments which are designed to demonstrate that the distinction is too vague; arguments which purport to show that the distinction is context-dependent (that the boundary which marks the distinction shifts over time as our theoretical knowledge and technological abilities evolve and progress); and arguments which are intended to show that the distinction cannot be made on the basis of an appeal to "privileged access".

The problem of vagueness I have already commented upon. So far as the argument which addresses the alleged "context-dependency" of the distinction is concerned, we may note that, even if sound, it does not necessarily pose a serious threat to GC. A distinction which is context-dependent in the sense indicated may serve as an adequate basis for a formulation of GC. The same may be said for the arguments directed against the possibility that the distinction may be made or justified by reference to something "given" in the sense advanced by radical empiricism or phenomenism. That the distinction cannot be explained in terms of considerations generated from a foundational epistemology of the sort represented by phenomenism is, I think, quite clear. Such an epistemology is, as I have previously urged, no longer considered viable. The basis of the distinction is therefore not a question of "privileged access". But again, the circumstance that the distinction is, in the last analysis, a pragmatic one (a matter of degree rather than one of kind) does not entail that it is neither an important nor a genuine one.

There is, however, another type of difficulty associated with the notion of empirical equivalence. In the case of complex theories, it may be extremely difficult to assess their relative empirical content and to establish their strict empirical equivalence. The reason for this is that

there is no general decision procedure for observational equivalence. This follows straightforwardly from the fact that there exists no decision procedure for theoremhood in an individual theory. There is a positive test for theoremhood (produce a derivation of the required conclusion), and hence for observational equivalence. But there is no negative test; failure to provide a derivation does not establish that none exists, since there is no decision procedure for theoremhood in a first-order language.<sup>15</sup>

In the case where two theories are strictly intertranslatable, we may have some assurance that they are empirically equivalent. But any pair of theories which exemplify GC are not isomorphic. This does not mean that GC is incoherent. Two theories certainly may be empirically equivalent. The difficulty lies with establishing their equivalence. There is thus a heavy burden on the part of the conventionalist who must persuade himself, and others, that two complex theories which he maintains exemplify GC are indeed strictly equivalent.

#### *Craig's Theorem*

Some words concerning the relevance of W. Craig's theorem (1956) to the present discussion might be deemed appropriate. I have already discussed Craig's work in Chapter 2. Recall that this result constitutes an algorithm for the re-axiomatization of a theory in a restricted vocabulary. Utilizing Craig's result, we may achieve the following. Given a standard scientific theory  $T$  expressed in a language  $L$  (which contains both an observational and a theoretical vocabulary), it will always be possible to construct another theory  $T^*$ , formulated in a language  $L^*$  (where  $L^*$  has the same logical apparatus as  $L$ , but where the descriptive terms of  $L^*$  will consist only of the observational vocabulary of  $T$ ), which has exactly the same observational consequences as the original theory  $T$ .  $T$  and  $T^*$  will thus be observationally equivalent, though they will not be isomorphic. In fact, while a standard theory will generally have only a small finite number of axioms, the "Craigian" counterpart to such a theory may have an infinite number.

Thus, Craig's Theorem presents us with the possibility of constructing two "theories" which are guaranteed to be empirically equivalent, and yet which are also known not to be intertranslatable. It does not, however, serve to undermine our contention that no sound

argument has ever been produced to support the thesis that alternative scientific theories may exist which are empirically equivalent, but genuinely distinct. In the first place, as we saw previously, the Craigian counterpart to a regular theory is indisputably not itself a "scientific" theory: that is, it fails some of the essential requirements of scientific discourse *per se*. Secondly, while it is true that a regular theory and its Craigian analogue are not intertranslatable -- although we have an effective procedure for generating  $T^*$  from  $T$ , we cannot reproduce  $T$  from  $T^*$  -- neither are they truly distinct. They are not "distinct" since  $T^*$  is generated by an effective mechanical procedure from  $T$  and since, indeed, this is the only way of acquiring  $T^*$ .

*The Lowenheim-Skolem Theorem*

We have just drawn attention to Craig's Theorem and discussed its possible implications for conventionalism. Before concluding, a second result which might appear to have some bearing on the issues at hand needs to be considered, namely, the Lowenheim-Skolem Theorem (Bell & Moshe, 1977, pp. 168-173). This theorem asserts that any satisfiable first-order theory expressed in a countable language has a countable model. Of particular interest is the strong form of the theorem which is referred to as the "downward" Lowenheim-Skolem Theorem. The downward version of the theorem states that any satisfiable first-order theory expressed in a countable language exhibits a countable model which is a submodel of a given (nondenumerable) model.

The result would seem, at first blush, to lend some support to conventionalism, or at least to open the way to a new formulation of the doctrine which might prove sound. It indicates that there may exist two models for a given theory and that, moreover, these models may be non-isomorphic. Furthermore, satisfaction of the various requirements encapsulated in our earlier formulation of GC -- including syntactic requirements such as austerity -- does not exempt a theory from this result. Consequently, it may occur to the diehard conventionalist that there remains the option of retreating entirely to semantic considerations.

So inspired, he may try to develop his position along the following lines:

"Conventionalism is a true and significant claim which is not crucially dependent on the syntax of our theories. It is a claim about the structures in which the disinterpreted calculus of an empirical theory (that component of an empirical theory which corresponds to the concept of a "theory" in the formal of model-theoretic sense of the term) is satisfied: its models.

Conventionalism is thus not a claim about the underdetermination of our theories in any sense which can be captured through essentially syntactical considerations. Rather, it is a claim about the underdetermination of the models for a given theory. It is the various possible models of a given theory -- and hence the ontologies conveyed by these models -- which are underdetermined." The conventionalist would then presumably assert that this conception of conventionalism derives support from the Lowenheim-Skolem Theorem.

It is the spectre of precisely this sort of "conventionalism" which is raised by Putnam in an article entitled "Models and Reality" (1980). Putnam addresses the problem of how we can specify which of the models that satisfy a theory is its intended model:

Now the arguments that Skolem gave, and that shows that the "intuitive notion of a set" (if there is such a thing) is not "captured" by any formal system, shows that even a *formalization of total science* (if one could construct such a thing), or even a *formalization of all our beliefs* (whether they count as "science" or not), could not rule out denumerable interpretations, and, *a fortiori* could not rule out unintended interpretations of this notion. (p. 466)

Putnam proceeds to outline a solution to this problem of specifying the intended interpretation of a first-order theory. Whether this solution is successful or not is an issue which is not really relevant to an assessment of the form of conventionalism we are presently contemplating. Let me explain this briefly. If we grant that the Lowenheim-Skolem Theorem really has implications regarding the possibility that the models of our theories (and hence the ontologies that they convey) may be underdetermined in a significant sense, then it is of no help to assert that we can specify the intended models. This only ensures that we can know exactly what we are saying, so to speak. It does not follow that what we are saying is necessarily true

An assurance that we can know exactly what we are asserting leaves intact the metaphysical and epistemological problems suggested by the existence of alternative ontologies which are compatible with the syntax of a given theory. Even if we can know which is the intended model, how can we be sure that this particular model, as opposed to any alternative "unintended model", reflects the true ontology of the world (assuming, as I do, that the latter phrase is meaningful)? We seem to be confronted with the circumstance that it may only be possible to know the "syntax" of the world, that the "semantics" may escape our theoretical nets: a modern analogue, perhaps, of the Aristotelian maxim that knowledge is of form, not of matter.

A closer look at the Lowenheim-Skolem Theorem will put to rest any fears that the conventionalist is building a case for himself. The theorem asserts that if we are presented with a nondenumerable model,  $M$ , of a theory  $T$ , which is formulated in a countable language, then we can always find another model  $M'$  of  $T$ , where  $M'$  is a countable model. But now comes the sticking point. The predicate symbols contained in  $M'$ , while restricted to a domain of suitable size, will represent the same relations as they did in  $M$ . Thus, the Lowenheim-Skolem Theorem only has implications regarding the size of the domains of our models. This hardly qualifies as the basis for an interesting or significant form of conventionalism. If all the conventionalist wishes to assert is that our theories, which may include "magnitudes" (real numbers) and therefore have nondenumerable models, also have countable models, then we may be willing to grant this point without feeling that *v.e* have conceded anything of real note. I would be willing to accept that the incorporation of the real number system in our theories is an instance of a *local* convention or definitional stipulation. (Theories do contain definitional elements, or local conventions, but the existence of such cannot be taken as a basis for asserting that theories as a whole are merely arbitrary stipulations.)

Furthermore, it may even be that the intended interpretations of our theories, even our physical theories, are actually denumerable *o.c.s.* One current cosmology posits the existence of a finite number of particles in the universe. Presumably, if once we produced a formulation of this cosmology which relied on the rational number system rather than the real number system,

then we would have a genuinely countable model of the universe. Of course, such a formulation would inevitably prove cumbersome and impractical from a computational standpoint, but we are not concerned with matters of mere convenience here. In fact if we begin from the standpoint that the "intended" model of the universe is just such a truly countable one, then the problem posed by the "Skolemization" of nondenumerable models does not necessarily arise.

What we have arrived at is the following situation. There is no pair of theories which exemplifies GC. Under careful analysis and with a more rigorous characterization, we see that GC is somewhat counterintuitive and there is no general logical grounds on which to build a case for the truth of GC. There remains the possibility that a pair of theories instantiating GC may arise, or that a formal existence proof may be forthcoming. But we are justified in doubting these possibilities. The burden of proof sits squarely on the shoulders of the conventionalist, and it has not been met. We can conclude, then, that the status of GC is not sufficient to build a case for relativism.

#### *A Final Argument for Relativism: Values as the Basis of Theory Selection*

David Thomas (1979) has contended that social sciences differ from the natural sciences insofar as values function as an additional criterion of theory selection in social science. He maintains that this is not only different from the situation in the natural sciences but also desirable, given the wide range of questions which must be addressed in social research and the limited scope of theories extant.

Thomas' arguments can be construed as supporting both a form of orthogonalism and relativism. The argument for orthogonalism flows from the premise that social science theories focus largely on different issues. This is different from orthogonalism as defined earlier in this chapter, where it was construed as a position concerning the relationship between qualitative and scientific traditions. But, if we accept the premise that theories of social science, even though sharing certain basic methodological precepts, can only be expected to deal with a limited number of dimensions of social phenomena, then a form of orthogonalism follows. Each successful theory may then serve as a model or paradigm for further developments in its specific

area of competence.

An argument for relativism can also be formulated from Thomas' position. To the extent that two theories may overlap in their areas of theoretical competence, Thomas believes that the choice between them is not primarily an epistemic one but rather one determined by one's values.

Let us consider the basic arguments Thomas musters. In the first place, he claims that it is an empirical fact that values play a decisive role in determining which of several theories will be selected. But this is hardly a conclusive point. We could decide that social scientists should not allow such considerations to influence choices, and proceed to train new researchers accordingly. In the absence of a good argument to establish that theoretical monism is desirable, the truth that social science is currently pluralistic does not further the case for pluralism. Certainly in the natural sciences pluralism is generally regarded as an earmark of undeveloped science. It is interesting to note that historically, in the case of less mature physics, values played a significant role in theory selection. There is no better example of this, perhaps, than the difficulties encountered in the transition to a heliocentric model of the universe.

Thomas' main argument against theoretical monism is predicated on the failings or limitations of Soviet sociology. This school of thought represents, he thinks, the closest thing to a true paradigm that is to be encountered in the social sciences. By "paradigm" he means, apparently, a truly dominant or monolithic approach to theorizing in social science. Soviet sociology does not appear to be any more conspicuously successful than alternative Western traditions.

Thomas argues that certain questions cannot even be raised within the Soviet tradition. For example, he asserts that its commitment to Marxism does not permit the study of antagonistic social classes because this notion is incompatible with historical materialism, the theoretical cornerstone of the Soviet school.

There is, however, a basic defect in Thomas' reasoning. Surely one's adherence to historical materialism as a theoretical position does not entail that one must be convinced that



Soviet society is a perfect instantiation of it. The fact that it is politically unacceptable to raise such questions does not mean that they cannot be framed or conceptualized. The limitations of Soviet sociology may thus be due, in certain respects, to political constraints rather than theoretical monism. In fact, these constraints appear to be in the process of being relaxed and it will be interesting to observe the effects of the current wave of political change on Soviet social science.

In brief, Thomas' main arguments for theoretical pluralism as a methodological precept for the social sciences are not convincing. Hence they do not provide any sound basis for relativism or orthogonalism.

#### *Summary of the Case for Relativism*

Overall, then, there seems to be no good case for relativism, despite its current fashionability. In particular, it cannot be based on either incommensurability or the thesis of the underdetermination of theory by empirical evidence, since the arguments in support of these philosophical positions are flawed and inadequate. It has been argued in this chapter that the weight of argument and presumption is actually against these theses. Nor, as we have just seen, can the circumstance that values influence theory choice justify relativism.

Ironically, Garrison's strategy is to accept, somewhat uncritically, the principles on which relativism is asserted, but then deny we need to capitulate to relativism. On his view we have only to expand the audience of those "competent to judge" to include practitioners. They have, he says, a common language and nonscientific ground on which to judge between competing scientific theories (Garrison, 1986, p. 18).

But this gambit will not do. If strong incommensurability, for example, is true, then the practitioners' common language is incommensurable with the various theories promulgated. There is no independent, theory-neutral language in which the results of theories may be expressed, compared and evaluated. The denial of such a language is the whole point of incommensurability. Having once assumed the mantle of incommensurability, we cannot proceed to sacrifice it on the altar of a theory-neutral or pretheoretic practitioner's language.

The sounder position, clearly, is to bring into question the warrant of the various principles on the basis of which a case for relativism might be constructed.

### *Conclusion*

Three dimensions of the relationship between the qualitative and quantitative have been assessed in this chapter: orthogonalism, compatibilism, and relativism. The major arguments for each have been presented and weighed, with the following conclusions.

First, there is a trend emerging to favour qualitative inquiry over scientific method in a variety of areas. One notable example of this trend was presented in the form of Neuman's polemic supporting a qualitative approach for the development and validation of precise prescriptions for the design of CAI. The significance of this particular article is enhanced by two circumstances. To begin with, it was published in a journal, *Educational Communications and Technology Journal*, which is central in the literature of educational technology in North America. Secondly, this is a publication which until very recently has been highly biased towards the scientific approach in terms of content. A review of the articles published in this journal over the last five years shows that approximately eighty percent are of the quantitative-objectivist variety. The remaining twenty percent includes articles that may be characterized as related to "issues" in the field or as constituting analytical or conceptual pieces. Thus, one may anticipate that Neuman's article may be influential, and one may also suspect that its publication reflects its alignment with current tendencies.

It was concluded, however, that while orthogonalism may be true as an empirical description of how qualitative methodology is being utilized in educational technology, the quantitative and qualitative approaches are more legitimately viewed as basically orthogonal. This follows from the premise that the goal of qualitative methodology is to capture and encapsulate what is unique about a situation or phenomenon, while the goal of the scientific approach is to determine what elements are due to regularities or uniformities which can be formulated in generalizations. Not all questions are amenable to the scientific approach, but proponents of the scientific approach have not always been too sensitive to this consideration.

The tendency to view qualitative and quantitative approaches as coextensive or non-orthogonal may owe a great deal, also, to the misapplication of the scientific approach to issues which do not admit of general theoretical explanations and which would be more appropriately addressed by some form of qualitative inquiry.

The arguments for compatibilism were also rejected. Several different levels of compatibilism were identified: compatibilism in terms of data, compatibilism in terms of propositions or knowledge claims, compatibilism in terms of procedures, and finally compatibilism in terms of basic paradigmatic assumptions about e.g., the nature of the objects of inquiry, and the nature of central notions such as "truth" and "validity". It was held that the level of paradigmatic assumptions was the crucial and interesting one and that compatibility at other levels could not obviate differences at the paradigmatic level. Thus, the *prima facie* case for incompatibilism was taken to stand, given the failure of the arguments for compatibilism

Most recently, Howe and Eisenhart (1990) have argued that the debate about standards for research, which has been played out over the last decades in terms of the quantitative-qualitative debate, should be conceptualized in terms of "logics in use", or the actual practices of researchers, rather than in terms of abstract epistemological considerations of the kind I include as constituting "assumptive frameworks" or "paradigmatic assumptions". Their argument essentially comes down to this: The philosophical or epistemological battle which has been fought has been waged along the axis of positivism versus anti-positivism. But since positivism has been thoroughly discounted in several respects, this debate is finished and the avenue for epistemological debate is closed. They write:

The common strategy of grounding qualitative research in an alternative paradigm creates a procrustean bed for itself by assuming that it must coexist with positivism. Refusing to entertain positivism as a viable epistemological doctrine -- a refusal that is now univocal within the philosophy of science -- is how to avoid this procrustean bed . . . Once positivism is removed from the scene, the positivist-alternative split, along with its various dualisms, collapses; the upshot is that standards must be anchored wholly

within a non-positivist perspective, which is to say they must be anchored nowhere other than in logics in use, in the judgements, purposes and values that make up research activities themselves. (p. 8)

This argument presupposes that the only epistemological battle to be fought must be defined along the positivist-anti-positivist line. But there are dimensions and dualisms which survive the collapse of positivism. In particular, what is left intact is the fundamental debate featuring the idealist assumptions of qualitative inquiry versus the realist assumptions of contemporary science (along with their respective implications for the concepts of "truth" and "validity"). It should be noted, too, that the extreme empiricism advocated by logical positivism is actually closer to the idealism of the qualitative approach than is the current, more moderate form empiricism has taken. Radical empiricism inevitably begets some form of idealism and solipsism: perhaps the best illustration of this relationship is the epistemology of Bishop Berkeley, the seventeenth century British empiricist and idealist. So the collapse of the positivists' position plainly does not dissolve all the epistemological differences between the assumptive framework of qualitative inquiry and a reformed philosophy of naturalistic social science, nor does it eliminate the significance of these differences from cognitive and methodological standpoints.

The case for relativism was also considered, in considerable detail. In particular, the two major supports for relativism – incommensurability and GC – were rejected. Incommensurability was judged to be demonstrably false (or worse, incoherent). The assessment of GC was that it had not been established by instantiation or by any general logical considerations and that it appeared, on closer inspection and more careful formulation, implausible.

The failure of the arguments for relativism and compatibilism has important implications for current tendencies towards hybridization and triangulation of qualitative and quantitative approaches in methodology. The epistemological underpinning or justification for these approaches is wanting. Even if this were not so, the pragmatic and "logics in use" justification of these practices fails, also. There is no established body of knowledge attributable to these innovations which can serve to legitimate them. In the absence of such demonstrable

successes, it is tempting to argue that these composite approaches simply result in a dilution of methodology. What is required, rather, is a more rigorous and systematic approach in which a unitary methodology is pushed to its limits, and thereby subjected to a genuine test. This is certainly arguable in the case of the scientific approach, which has never been properly implemented in the social sciences.

In concluding this chapter, a few comments regarding the relationship among the issues of orthogonalism, compatibilism and relativism are in order. They are logically distinct. For example, if non-orthogonalism is accepted, one is still forced to decide the issue of compatibilism. If orthogonalism were held true (particularly if the sets of issues addressed by the two approaches were viewed as completely disjoint), then the issue of compatibilism could still be raised, although it would be largely irrelevant. Relativism, under this condition, would clearly be a non-issue. If orthogonalism is strictly true, then one's choice of method is determined by one's questions.

Compatibilism combined with non-orthogonalism might seem to furnish some support for relativism. But the relationship is by no means an entailment. There is still a possibility of constructing an argument to establish the superiority of one approach over the other, on cognitive grounds. Compatibilism does not even constitute a necessary condition for relativism, since the latter could conceivably be consistent with incompatibilism and non-orthogonalism.

## CHAPTER 6

### Concluding Remarks

I have outlined the basic arguments against the naturalistic approach in social/behavioural sciences and have concluded that none is decisive. Generally speaking, the arguments which have been mustered against naturalism in the literature of education and educational technology are not sound. Opposition to the naturalistic approach is motivated to some degree by its failure to deliver theoretical insights and to contribute to the solution of pressing educational problems. However, there is also an element of ideological opposition. The scientific approach is viewed by many critics as dehumanizing, and it is frequently associated with the *status quo* in education rather than with forces of change.

There is something to these criticisms. There may be an element of truth to the assertion that, historically, the scientific paradigm has not been linked with attempts to improve education and redress inequities through the radical reform and restructuring of educational institutions and processes. But this is not a criticism of the power of the paradigm as a mode of inquiry *per se*. It is also true, however, that educational research which follows the established pattern of the scientific approach, with its reliance on the statistical methods advocated by Fisher, tends to address the mean or the average. The theories and hypotheses developed and confirmed in this mould address the norm. Behaviour which falls outside one or two standard deviations remains unexplained. As a concomitant, there is an unfortunate tendency on the part of researchers to believe that different theories are required to explain the behaviour of different subsamples, rather than to search for more sophisticated, more inclusive theories.

This is one serious problem associated with the present naturalistic paradigm. There are, as we have seen, additional ones associated with the conception of theory and the techniques of theory testing and theory elaboration which form the bases of the "scientific" model. I have argued the limitations of these on epistemological and logical grounds and I have illustrated their pervasive influence in the research literature of educational technology. Unless we develop a more sophisticated appreciation of the characteristics of theoretical discourse, and adopt a more powerful approach to hypothesis testing, there is little hope that efforts undertaken

under the banner of "science" in the social/behavioural domain will lead to significant insights or anything approaching a theoretical base which might inform the development of a true "technology" of educational or instructional processes, if such a thing is indeed possible.

Unfortunately, the debate concerning the merits and potential of the scientific approach has not been conducted under conditions which would render it conducive to promoting the kinds of changes which must be contemplated. The terms of the debate have been set, for a considerable time, by external critics, especially proponents of qualitative modes of inquiry. The attacks launched by such critics often miss the mark. The epistemological arguments which are advanced are often irrelevant, in so far as they are aimed against a model of science based on operationism and logical positivism which naturalists should eschew, also. A more productive debate would focus on ways in which the scientific paradigm in the social and behavioural domains might be reformed and strengthened. The question whether such reforms are possible would be a legitimate one; but this question has hardly ever been posed within the qualitative-versus-quantitative forum, which seems to regard the existing naturalistic model as an unalterable given.

By allowing external critics to set the agenda for methodological debate two unfortunate trends have resulted. In the first place, proponents of the scientific approach have been distracted from pursuing methodological issues in a fruitful manner. And, secondly, qualitative inquirers have been so preoccupied with the job of establishing their own traditions by discrediting the established scientific approach, that they, too, have devoted comparatively little energy to dealing with pressing internal methodological issues of their own. It is distressing to note that there is an almost perfect symmetry between qualitative and quantitative inquiry in educational research as regards the improper implementation of their respective models or paradigms. Naturalists blithely ignore null results, explaining them away on *ad hoc* grounds; they accept all sorts of violations of the statistical models employed; they push the levels of data they are working with, using inferential statistics which require interval data with ordinal data; they place no value on the replicative studies their logic of hypothesis testing requires; they

make causal inferences on the basis of correlations. Qualitative inquirers, for their part, run quick and dirty "ethnographic" studies that are based on "thin" description. Sometimes the results are collected from multiple sites by different (possibly untrained) investigators using standard, predefined instruments over very short periods of time, and then synthesized by other members of the team in order to draw conclusions. Without prolonged engagement, thick description, persistent observation, peer debriefing, evaluation and fine-tuning of instrumentation, the creation of various audit trails and the use of other techniques designed to enhance the trustworthiness of qualitative inquiry, such studies cannot claim the status of "disciplined inquiry" (Cronbach & Suppes, 1969).

More recently, there has been a shift in the terms of the methodological debate, as described in the previous chapters. The trend is to ignore epistemological dimensions relating to the assessment of modes of inquiry. The justification offered for this new approach consists in either of two lines of reasoning.

The first is an appeal to pragmatism and "logics in use". The argument here is that the epistemological debate must focus on what researchers are actually doing. This tactic fails because there is simply little justification for what researchers are currently doing: there are no established, and no clearly emerging, traditions of inquiry which are establishing substantial bodies of systematic knowledge.

The second is an appeal to elements of post-positivist philosophy. The rejection of popular themes of logical empiricism, such as the fact-value dichotomy and the distinction between observation and theory, are taken to establish that no boundary can be drawn between naturalistic and qualitative paradigms, that no room for epistemological debate exists. Unfortunately, as we seen, the fact-value and theory-observation bifurcations are not quite so dead as critics of naturalism would like to suppose. Moreover, there remain other dimensions to the contrast between the qualitative-interpretivist and quantitative-objectivist paradigms: in particular, the dimension created by the opposing assumptions of realism and idealism.

The most pernicious aspect of the uncritical appeal to certain post-positivist principles



(some of which are no longer considered viable themselves) is that they provide a rationale for an extreme form of methodological anarchism. Rather than strengthening our methodologies (both the qualitative and quantitative varieties), we are watering them down in various ways. We are following a path away from disciplined inquiry. There is no virtue in being unmethodical if it merely leads to the most efficient confirmation of our prejudices, regardless how comforting that may be.

The current climate thus favours methodological pluralism. Qualitative research is viewed by many as a legitimate alternative to the scientific approach, even where the goal of inquiry is general, systematic knowledge. Others favour a blending of the approaches, and justify this with an appeal to relativism and compatibilism. Consequently, much of this dissertation has been devoted to examining the logical and epistemological bases for relativism and compatibilism. In the course of this examination I presented original arguments against the philosophical principles implicated in both.

The principal bases for relativism were the thesis of the underdetermination of theory by evidence (or global conventionalism) and the Kuhn-Feyerabend incommensurability thesis. Incommensurability rests on the meaning-variance thesis which states that terms acquire their meaning within a theory rather than within a language (in the informal senses of these terms). At best the meaning variance thesis is false, but the situation is probably worse than that; an argument can be constructed to show that it is incoherent.

Global conventionalism comes off only slightly better. Following a rigorous account of the meaning of global conventionalism it became clear that no pair of theories instantiating the thesis of conventionalism exists, that no general logical grounds exist to support the doctrine and that, further, the thesis is implausible when viewed from the vantage point of a clear formulation of what it entails.

The arguments for compatibilism turn largely on the rejection of the fact-value and theory-observation distinctions, distinctions which were upheld in this present work. Other elements that were identified as contributing to the acceptance of compatibilism included the

tendencies to draw parallels between procedures in qualitative inquiry and quantitative inquiry and to use the terms "validity" and "truth" as if they held the same meaning across the two paradigms. There was also some confusion evident concerning at what level the question of compatibilism should be raised: data, procedures, knowledge claims or logics of justification.

In general, the tone of this dissertation has been quite pessimistic. While the possibility of a viable naturalistic approach to research in educational technology has been defended, it has also been argued that the present naturalistic model suffers major defects and that consequently it is unlikely, even if fully and consistently implemented, to yield major advances in understanding and prediction. There are, however, some promising developments and directions in current research related to educational technology which should be duly noted.

It is encouraging to see that cognitive science, cognitive psychology and information processing approaches to understanding learning are beginning to influence and shape instructional systems design, and at least to supplement the behavioural psychology orientation that has influenced much of the history of educational technology. Behaviourism provided the basic principles concerning the "shaping" of behaviour and the effects of feedback which were the bases of earlier innovations such as programmed learning and PSI (personalized self-instruction). Recently, more detailed theories concerning memory processes and the evolution of understanding through the elaboration of our mental models of the world ("schemata") have begun to play a role in ISD and instructional theory (Di Vesta & Rieber, 1987; Bruning, 1983; Kember & Murphy, 1990; Tennyson, 1990; Merrill, Li & Jones, 1990a; 1990b; 1990c; Winn, 1990).

Glaser (1990) argues that a new emphasis on learning theory is emerging, following a long period during which the focus has been explicitly on performance. The performance orientation has been reflected not only in research initiatives but also in the cult of behavioural objectives (the premise that all learning outcomes should be specified in terms of observable behaviours) which has dominated instructional design and infiltrated curriculum development.

Glaser argues persuasively that learning theory has made a comeback. The context in which it has been revived is furnished by "studies that take principled approaches to the design

of instruction for complex forms of knowledge and skill" (p. 29). What is common to these studies is their grounding in explicit cognitive task analysis and, for the most part, the fact that they address narrow, well-defined knowledge domains and forms of competence. Glaser's review of these programs includes domains such as medical diagnosis, electronics troubleshooting, reading comprehension, computer programming, medical diagnosis, algebraic equations, and geometric proofs. He argues that the application of cognitive analysis to instructional design in these domains is leading up to the formulation of theoretical accounts of three forms of learning: the acquisition of the highly automatized, proceduralized knowledge which characterizes expert performance in problem-solving in specific domains; the development and application of metacognitive skills (internalized self-regulatory performance control strategies for knowledge acquisition), and; the structuring of knowledge in the form of mental models that enable problem-solving.

Instructional programs designed to foster the development of these kinds of learning operationalize key terms and provide a valuable test bed for the theory on which they are predicated. This assertion requires some comment, in light of our earlier rejection of Reigeluth's call for "formative evaluation" research. Recall that I argued that the logical relationship between instructional theory or instructional design and the products we develop was too weak for the success or failure of a designed product to confirm or disconfirm the model or theory on which it was based. The difference is that conventional procedural ID models and instructional "theory" based on behavioural principles are quite vague, while the cognitive theories which are implicated in the research to which Glaser refers are much more detailed and precise, and offer correspondingly tighter prescriptions. Therefore, the link between theory and product is tighter in the situations Glaser describes and the implications of the performance of the products for the validity of the theories on which they are grounded is correspondingly greater.

Moreover, many of the programs discussed by Glaser are implemented in the form of ITS technology (computer-based intelligent tutoring systems) (Mandl & Lesgold, 1988; Polson & Richardson, 1988; Wenger, 1987). ITS technology is distinguished from traditional or frame-

based forms of computer-based instruction by three characteristics: (1) An ITS includes an explicit representation of the knowledge domain. More precisely, it contains a representation of the expert's mental model of the domain and his inference strategies, captured in terms that are similar to how the expert represents these to himself. This is accomplished using expert systems technology, characteristically a collection of "production rules" of the "If . . . then . . ." variety. (2) Also included is an explicit model of the learner. Typically this comprises a "bug catalogue" of possible misconceptions that is elaborated and validated through empirical research with the target population. (3) Finally, an ITS includes an explicit pedagogical theory linking instructional strategies with learners' performance.

The requirements of creating executable programs necessitate a high degree of explicitness in the specification of expert's mental models, learner's mental models, and pedagogical theory, thus providing a solid test of the underlying theory.

Unfortunately, the high cost of ITS development limits this kind of work. ITS systems are custom tailored for specific domains and there are no general ITS development tools of the kind we encounter in traditional CAI, in the form of various shells and authoring languages. Moreover, much ITS development work has been carried on outside of educational technology, in the fields of engineering and artificial intelligence studies in computer science. It remains to be seen whether educational technology can become substantially involved in this type of work. The survival of the field may ultimately be tied to this question.

It should be remarked that the direction of ITS research and development is appealing from several perspectives. Not only does it hold promise for furthering our understanding of cognitive processes, but it also offers a counterbalance to the present prevailing tendencies of educational technology. Educational technology is concerned very much, at present, with the mass production of instructional artifacts in a cost effective manner. This is accomplished utilizing an industrial mode of production that exploits conventional ID processes and generic CAI authoring tools, and focuses on performance measures. ITS work forces us to come to grips with the qualitative aspects of learning and understanding. Learning cannot be conceived

merely as the acquisition of so many more units of the same performance measure; to learn is to acquire new schemata, new mental models and capacities, and these are the substance of ITS development.

Moreover, conventional educational technology focuses on learning to the exclusion of the consideration of teaching processes. The field emphasizes instructional and environmental variables which are viewed as more reliable and operationalizable than variables which describe teaching behaviours. Indeed, the standard educational technology approach is to replace teaching (which implies a certain degree of "on-line" or spontaneous analysis and response to instructional events) with pre-packaged instruction. This orientation is also reflected in the dearth of references to the burgeoning body of research concerning teaching processes and variables in the articles published in the literature of educational technology. ITS, however, makes pedagogical theory an issue in the realm of instructional technology, once more. Indeed, perhaps the most sophisticated and most promising work in pedagogical theorizing to date has been the careful analysis of tutoring methods that has been conducted in conjunction with ITS projects such as WEST and SPIRIT (Wenger, 1987, pp. 140-152).

However, these salutary developments I have been discussing do not mitigate the basic conclusions arrived at in this dissertation. To reiterate one last time: The debate concerning the relative merits of qualitative and quantitative inquiry is clouded by misconceptions. The quantitative-objectivist paradigm, though it suffers real defects, is being rejected outright or further diluted and degraded through hybridization, for the wrong reasons. Qualitative inquiry itself suffers major problems, especially regarding the issues of validity, reliability and generalizability, problems that are only exacerbated by the current tendency to focus on mere procedures in data collection, synthesis and analysis. This makes the abandonment of scientific method in favour of a qualitative paradigm less attractive than current fashion suggests. Methodological triangulation, another popular innovation, has no logical basis or rationale, given the rejection of compatibilism as an account of the relationship between the qualitative and quantitative paradigms. Hence, it cannot be seen as a solution to the problems currently

associated with each camp.

In general, the trend in research (and evaluation) seems to be towards a sort of methodological anarchism, to borrow a phrase from Feyerabend (1975). In one sense, the existence of a plurality of methodologies might be a good thing. In the end those which were most successful at developing a body of significant generalizations would survive, while the failures would eventually fall by the wayside. But this vision of competing methodologies is one in which well-defined, methodical approaches compete against one another. We seem, rather, to be in danger of abandoning disciplined inquiry, of repudiating method *per se*. The result, ultimately, may be that applied social science will be able to do no better than affirm the prejudices of those who conduct research.

## FOOTNOTES

1. The arguments against the radical empiricist strategy of trying to justify all knowledge claims by grounding them on incorrigible reports of pure sensory experience are not, of course, entirely independent of the arguments against the theoretical-observational distinction. If all language is "theory-laden", if there is no such thing as a pure observational language, then it follows that the language required for the positivist's foundational epistemology is simply not available. But, at the same time, it is possible to reject the positivists' solution to the problem of knowledge without abandoning the basic observational-theoretical distinction. One can simply reject foundationalism by refusing to accept that *any* statements are incorrigible, or by refusing to allow that the whole edifice of human knowledge could possibly be derived from sentences of the nature of the positivists' protocols.

2. It was not always so. Logical positivism spawned a Unity of Science Movement which posited the reducibility of all sciences, including psychology, to the laws and theories of physics. The history of science cuts against the grain of this vision, however. The various compartments of science have become increasingly independent, and this is a continuing trend: physical chemistry, for example, is farther removed from the branches of physics today than it was twenty years ago.

3. This term, "rationalist", supercedes Guba's previous use of the label "scientism" for the same purpose. In a footnote Guba admits unabashedly that the switch in terminology was necessitated by accusations that he had constructed a straw man. Interestingly, rather than confront the charge head on Guba has chosen to follow a more oblique route, apparently seeking to confound his enemies by hiding his straw man under a new name. This is by no means a new ploy -- but the accompanying forthright confession to what amounts to a rather tawdry rhetorical trick is certainly a novel twist.

4. There are further indications of this in the body of the paper. For example, consider the "persuasive" (Weldon, 1953) use of language pregnant with connotations in describing the naturalistic approach: "Practitioners of scientific inquiry, in the hard but especially in the soft sciences, often continue to act as if the (naturalistic) paradigm had validity, continuing to accept a position that is essentially analytic, reductionist, empiricist, associationist, reactivist, nomological, and monistic" (p. 235). Apparently, this is to be read as a litany of sins.

5. And thus there seems to be a strong connection between theoreticity in this sense and theoreticity in the other sense, since discourse which is highly theoretical in the second sense is also invariably discourse which ranks highest in explanatory power.

6. This particular argument against naturalism raises the following consideration. Critics of the naturalistic approach, in the guise in which it appears in the social/behavioural sciences, often complain that it "does not reflect emergent (recent) philosophies of science". One of the

principle tenets of the "emergent" philosophies they are alluding to is that observation is unavoidably coloured by theoretical assumptions or hypotheses. Indeed, the argument that is made by these philosophies is that theory determines in large part what will be selected as data, how that data will be interpreted, what will count as evidence. Anti-naturalists have a tendency to appeal to these philosophical principles when it suits their needs to develop a case against the naturalistic approach, but then to assume the possibility of neutral investigation when they desire to argue the superiority of qualitative approaches. More recently, in the anthropological literature, ethnographers have themselves become more preoccupied with the problem of the assumptions and prejudices which must inevitably influence the interpretations and procedures of ethnographic inquiry (McClancy, 1986). In the educational literature, qualitative inquiry is still in a missionary mode, it seems, and has not yet reached the point of entrenchment and maturity which might encourage such self-critical reflection (Rist, 1980).

7. Of course, General Systems Theory does assert this, but it then proceeds with a methodology -- the search for isomorphisms among existing empirical theories which might be taken as evidence of their issuance from a single theory of general systems -- that is itself entirely empirical.

8 What is particularly annoying, however, is that while the term "logical positivism" is thrown about with great abandon few players in the game seem to know what it really means. Karl Popper is routinely described as a logical positivist, though he is perhaps the most important critic of the movement. Dr. Guba, one of the most outspoken critics of the scientific approach, describes logical positivism in a footnote of his 1982 paper (Guba & Lincoln, 1982) as a nineteenth-century doctrine. For someone so concerned about this school of thought this is shoddy scholarship. Logical positivism is in fact the name attributed to the views advanced by a group of thinkers based chiefly in Vienna and initiated during the 1930's.

9. The argument that hybridization leads to a dilution of scientific method and hence potentially a weaker mode of inquiry might be possible even if the view I will label "compatibilism" were secured.

10. Howe is certainly not alone in the educational field in asserting this view; several other commentators defer specifically to his thinking on this point (cf. Guba, 1979; Guba & Lincoln, 1982; Gage, 1989; Tranel, 1981).

11. The difference between Quine's brand of pragmatism, as advanced in "Two Dogmas of Empiricism" (1980), and Duhem's thesis seems to be largely simply a question of breadth. Quine holds that any statement, S, can be maintained, whatsoever -- providing compensating adjustments are made to the interpretation (truth values) of the other statements which, together with S, constitute a consistent system of knowledge and belief. Quine explicitly allows that these adjustments may have to include the laws of logic themselves.

Duhem, on the other hand, only maintains that there is a certain degree of slack in the testing of scientific theories. This "slack", as explained, is afforded by the roles played by a number of elements including cognate theory, assumptions about the functioning of instrumentation, statements of initial conditions and rules of correspondence. Given sufficient slack, a favoured experimental or theoretical hypothesis may be saved in the face of recalcitrant observations by making compensatory adjustments with respect to these additional elements.



Duhem does not attempt to extend his thesis to involve potentially the rules of logic or to encompass our common sense beliefs concerning the world.

It may appear, then, that Quine's position is simply a logical extension of Duhem's. But this view of the relationship between the two obscures an important difference. Duhem's position is based squarely on deductive logic -- in particular the *modus tollens* rule. Quine's, on the other hand, is supported by a particular view of language and a particular epistemology. Quine's position requires that one accept that any consistent, sufficiently rich, language (set of statements) has an interpretation which is a model of the world. Duhem's less extreme thesis does not require such an assumption.

12. Suppes (1960) suggests that the question of the reducibility of one branch of science to another can be restated formally as the demand for an appropriate representation theorem. A theory  $T_1$  is reducible to a theory  $T_2$  if, and only if, for any submodel of  $T_1$ , there exists an isomorphic model within  $T_2$ .

13. It should be emphasized that this is merely an alternative identified by Glymour. He does not defend this criterion in this particular article.

14. The verificationist principle or criterion of meaning was a central plank in the logical empiricists platform. Briefly stated, the criterion asserts that a sentence is meaningful if and only if one can specify the procedures by which it might be tested or confirmed. Thus, the principle excludes all metaphysical propositions. A more radical thesis to the effect that the meaning of a proposition is equivalent to, or reducible to, the procedures for testing it, was also advanced, and was even more problematic. (The difference between these two theses is the difference between a criterion or test that identifies which statements are meaningful, and a theory of meaning which specifies what is the meaning of a meaningful statement.) There are close affinities between the verificationist theory of meaning and the operationalist position described in Chapter 3. Both are expressions of radical empiricism.

15. There is no decision procedure for observational equivalence. Suppose that we have two theories,  $T_1$  and  $T_2$ . Take a sentence  $S$  in the observation language which is known to be a theorem of  $T_1$ . If we can produce a derivation of  $S$  from  $T_2$ , we will have confirmed the equivalence of  $T_1$  and  $T_2$  as regards  $S$ . If, on the other hand, we can derive  $\sim S$  from  $T_2$ , then we will have established that the two theories are not empirically equivalent. But what happens if neither  $S$  nor its negation is forthcoming from  $T_2$ ? We are faced with a dilemma in that case. It may be that  $T_2$  is simply not complete with respect to the observational language (i.e.,  $T_2$  features neither  $S$  or  $\sim S$  as a theorem), or it may be that either  $S$  or  $\sim S$  is indeed a theorem of  $T_2$  and we have simply not succeeded in formulating the required derivation. This is just to say that we lack a general decision procedure for observational equivalence. Another problem arises from the fact that complex theories will generally yield an infinite number of predictions (some of which for practical reasons may not be testable) so that it may not be possible to actually compare exhaustively their respective empirical content.

## REFERENCES

- Adams, E. W. (1959). The foundations of rigid body mechanics and the derivation of its laws from those of particle mechanics. In L. Henkin, P. Suppes and A. Tarski (Eds.), *The axiomatic method*. Amsterdam: North-Holland Publishing Co.
- AECT. (1972). The field of educational technology: A statement of definition. *Audiovisual Instruction*, 36-44.
- Anderson, G. L. (1989). Critical ethnography in education: Origins, current status and new directions. *Review of Educational Research*, 59(3), 249-270.
- Angel, R. B. (1980). *Relativity: The theory and its philosophy*. Oxford: Pergamon Press.
- Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development*, 32(4), 2-16.
- Angeles, P. A. (1981). *Dictionary of philosophy*. New York: Harper & Row.
- Angus, L. B. (1986). Developments in ethnographic research in education: From interpretive to critical ethnography. *Journal of Research and Development in Education*, 20(1), 59-67.
- Asher, H. B. (1983). *Causal modeling*. Beverly Hills, CA: Sage Publications.
- Ausubel, D. P. (1968). In defence of advance organizers: a reply to the critics. *Review of Educational Research*, 48(2), 251-257.
- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune and Stratton.
- Baath, J. (1981). On the nature of distance education. *Distance Education*, 2(2), 212-219.
- Barker, S. F. (1957). *Induction and hypotheses*. New York: Ithica, NY: Cornell University Press.
- Becker, H. J. (1987). The importance of a methodology that maximizes falsifiability: Its applicability to research about Logo. *Educational Researcher*, 16, 11-16.
- Beckwith, D. (1989). Fashioning the future educational technology: An invitation. *Canadian Journal of Educational Communication*, 18(2), 127-131.
- Bell, J. & Moshe, M. (1977). *A course in mathematical logic*. Amsterdam: North Holland Publishing Company.
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modelling. *Annual Review of Psychology*, 31, 419-456.
- Bergman, G. (1953). Theoretical psychology. *Annual Review of Psychology*, 4, 435-458.
- Bernard, R., Naidu, S. & Amundsen, C. (1990). Choosing instructional variables for practice and research in distance education. In press.
- Berry, M. V. (1977). Technology. In A. Bullock & O. Stallybrass (Eds.), *The fontana dictionary of modern thought* (p. 625). London: Collins.

- Beth, E. (1961). Semantics of physical theories. In H. Freudenthal (Ed.), *The concept and the role of the model in mathematics and natural and social sciences* (pp. 48-51). Dordrecht: Reidel.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw Hill
- Bloom, B. S. (1956). *Taxonomy of educational objectives, Handbook 1: Cognitive domain*. New York: David McKay Inc.
- Borg, W. R. & Gall, M. D. (1989). *Educational research: An introduction*. New York: Longman
- Borg, W. R. & Gall, M. D. (1983). *Educational research: An introduction*. New York: Longman
- Brahm, M. L. (1973). The grounding of the technologist. In *Aspects of educational technology VII*. pp. 69-77.
- Bratton, B. (1988). Roles for educational technologists by the year 2000. *Journal of Instructional Development*, 11(3), 7-9.
- Bridgman, P. W. (1927). *The logic of modern physics*. New York: MacMillan.
- Briggs, L. J. (1977). *Instructional design: Principles and applications*. Englewood Cliffs, NJ: Educational Technology Publications.
- Bruner, J. S. (1971). *Towards a theory of instruction*. Cambridge, Mass.: Harvard University Press.
- Bruner, J. S., Goodnow, J. J. and Austin, G. A. (1956). *A study of thinking*. New York: Wiley.
- Bruning, I. L. (1983). An information processing approach to a theory of instruction. *Educational Communications Technology Journal*, 31(2), 91-101.
- Bunge, M. (1972). A program for the semantics of science. *Journal of Philosophical Logic*, 1, 317-328.
- Bunge, M. (1963). *The myth of simplicity*. Englewood Cliffs, NJ: Prentice Hall.
- Bunge, M. (1962). The complexity of simplicity. *Journal of Philosophy*, 59(5), 113-135.
- Carpenter, C. R. (1953). A theoretical orientation for instructional film research. *AV Communications Review*, 1, 38-52.
- Carr, W. (1983). Can educational research be scientific? *Journal of Philosophy of Education*, 17(1), pp. 35-43.
- Cassidy, M. F., & Knowlton, H. Q. (1983). Visual literacy: A failed metaphor? *Educational Communications and Technology Journal*, 31(2), 67-90.
- Clark, R. E. (1989). Current progress and future directions for research in instructional technology. *Educational Technology, Research and Development*, 37(1), 57-66.
- Clark, R. E. (1984). "Where's the beef?": A reply to Heinich. *Educational Communications and Technology Journal*, 32(4), 229-232.
- Collin, F. (1985). *Theory and understanding: A critique of interpretive social science*. Oxford: Basil Blackwell Ltd.
- Craig, W. (1956). Replacement of auxiliary expressions. *Philosophical Review*, 65, 38-55

- Cronbach, L. J. (1988). Playing with chaos. *Educational Researcher*, 17, 46-49.
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30(2), 116-127.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12(11), 671-684.
- Cronbach, L. J. & Suppes, P. (Eds.). (1969). *Research for tomorrow's schools: Disciplined inquiry for education*. New York: MacMillan.
- Dale, F. (1946). *Audio-visual methods in teaching*. New York: Dryden Press.
- Dar, R. (1987). Another look at Meehl, Lakatos, and the scientific practices of psychologists. *American Psychologist*, 42(2), 145-151.
- Davidson, D. (1971). Agency. In R. Blinkley, R. Bronough & A. Marras (Eds.), *Agent, action and reason* (pp. 3-25). Toronto: University of Toronto Press.
- Davidson, D. (1966). The logical form of action sentences. In N. Rescher (Ed.), *The logic of decision and action*. Pittsburgh: University of Pittsburgh, 81-95.
- Davidson, D. (1964). Causal relations. *Journal of Philosophy*, 64, 691-703.
- Davidson, D. (1963). Actions, reasons and causes. *Journal of Philosophy*, 60, 685-700.
- Davies, I. K. (1978). Prologue: Educational technology: Archetypes, paradigms and models. In J. Hartley & I. K. Davies (Eds.), *Contributions to an educational technology, volume 2* (pp. 2-14). London: Kogan Page.
- de Brisson, C. (1987). *Understanding technological change*. Montreal: Black Rose Books.
- Di Vesta, F. & Rieber, L. P. (1987). Characteristics of cognitive engineering: The next generation of instructional systems. *Educational Communications and Technology Journal*, 35(4),
- Dick, W. (1981). Instructional design models: Future trends and issues. *Educational Technology*, 21(7), 29-32.
- Donmoyer, R. (1985). The rescue from relativism: Two failed attempts and an alternative strategy. *Educational Researcher*, 14, 13-20.
- Duchastel, P. (1989). The upcoming of learning technology. *Canadian Journal of Educational Communication*, 18(2), 137-139.
- Duhem, P. (1974). *The aim and structure of physical theory* (P. P. Weiner, Trans.). New York: Anthem.
- Dyer, J. B. (1979). *Understanding and evaluating educational research*. New York: Addison Wesley.
- Egan, K. (1988). The analytic and the arbitrary in education research. *Canadian Journal of Education*, 13(1), 69-81.
- Ellul, J. (1964). *The technological society*. New York: Alfred A. Knopf.

- Estes, W. K. & Suppes, P. Foundations of linear models. In R. R. Bush and W. K. Estes (Eds.), *Studies in mathematical learning theory*, (Chapter 8). Stanford, CA: Stanford University Press.
- Fetterman, D. M. (1982). Ethnography in educational research: The dynamics of diffusion. *Educational Researcher*, 11, 17-22.
- Feyerabend, P. K. (1975). *Against method*. London: NLB
- Finn, J. O. (1953). Professionalizing the audio-visual field. *AV Communication Review*, 1, 6-17.
- Follesdal, D. (1979). Hermeneutics and the hypothetico-deductive method. *Dialectica*, 33(3-4), 319-336.
- Friedman, M. (1977). Simultaneity in Newtonian mechanics and Special Relativity. In J. Earman, C. N. Glymour & J. J. Stachel (Eds.), *Foundations of space-time theories, Minnesota Studies in the Philosophy of Science*, vol 8.
- Friedman, M. (1972a). Grunbaum on the conventionality of geometry. *Synthese*, 24, 219-235.
- Friedman, M. (1972b). Empirical simplicity as testability. *British Journal for the Philosophy of Science*, 23, 25-33.
- Gage, N. L. (1989). The paradigm wars and their aftermath: A "historical" sketch of research on teaching since 1989. *Educational Researcher*, 18, 4-10.
- Gardner, M. (1976). The unintelligibility of observational equivalence. *Proceedings of the 1976 biennial meeting of the Philosophy of Science Association*, 1, 104-116.
- Garrison, J. W. (1986). Some principles of postpositivistic science. *Educational Researcher*, 15, 12-18.
- Gibson, J. J. (1954). A theory of pictorial perception. *AV Communications Review*, 2, 2-23.
- Glaser, R. (1990). The reemergence of learning theory within instructional research. *American Psychologist*, 45(1), 29-37.
- Glymour, C. (1977). The epistemology of geometry. *Nous*, 11, 229-251.
- Glymour, C. (1970). Theoretical realism and theoretical equivalence. *Boston studies in the philosophy of science*, Vol VIII. 275-287.
- Good, T. L. & Brophy, J. E. (1986). *Educational psychology (3rd ed.)*. New York: Longman.
- Goodman, N. (1968). *Languages of art*. Indianapolis: Bobbs-Merrill.
- Grunbaum, A. (1960). The Duhem argument. *Philosophy of Science*, 27(1), 75-87.
- Guba, E. G., & Lincoln, Y. S. (1982). Epistemological and methodological bases of naturalistic inquiry. *Educational Communications and Technology Journal*, 30(4), 233-252.
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communications and Technology Journal*, 29, 75-91.
- Guba, E. G. (1979). Naturalistic inquiry. *Improving Human Performance Quarterly*, 8(4), 268-276.

- Gustafson, K. & Bratton, B. (1984). Instructional improvement centers in higher education: A status report. *Journal of Instructional Development*, 7(2), 2-7.
- Hanafin, M. J. (1989). The death of educational technology has been greatly exaggerated. *Canadian Journal of Educational Communication*, 18(2), 140-142.
- Hanson, N. R. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Hare, R. M. (1963). *Freedom and reason*. Oxford: Oxford University Press.
- Hayes, R. T. & Singer, M. J. (1989). *Simulation fidelity in training systems design: Bridging the gap between reality and training*. Amsterdam: Springer Verlag.
- Heinich, R. (1984). The proper study of instructional technology. *Educational Communications and Technology Journal*, 32(2), 67-87.
- Herriot, R. E., & Firestone, W. A. (1983). Multisite qualitative policy research: Optimizing description and generalizability. *Educational Researcher*, 12, 14-19.
- Hess, M. (1966). *Models and analogies in science*. Notre Dame, Indiana: University of Notre Dame Press.
- Hess, M. (1963). A new look at scientific explanation. *Review of Metaphysics*, 17, 98-108.
- Holmberg, B. (1986). A discipline of distance education. *Journal of Distance Education*, 1(1), 25-40.
- Holmberg, B. (1985). Theoretical underpinnings of mediated teaching as self-contained instruction -- a suggested theory. *Programmed Learning and Educational Technology*, 22(1), 16-25.
- Holmberg, B. (1983). Distance education as an academic discipline. *International Council for Distance Education Bulletin*, 2, 50-53.
- Hooker, C. A. (1968). Craigian transformation. *American Philosophical Quarterly*, 5, 152-163.
- Hooper, R. (1971). A framework for studying instructional technology. In S. G. Tichton & R. R. Bowker (Eds.), *To improve learning: An evaluation of educational technology, vol II*. New York.
- Horwich, P. (1982). How to choose between empirically indistinguishable theories. *Journal of Philosophy*, 79(2), 62-77.
- House, D. & Bratton, B. (1986). *An evaluation of longitudinal curriculum changes in instructional design and educational psychology doctoral programs*. Paper presented at the Association for Educational Communications and Technology Conference, Las Vegas, NV.
- Howe, K. R. (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, 17, 10-16.
- Howe, K. R. (1985). Two dogmas of educational research. *Educational Researcher*, 14, 10-18.
- Howe, K., & Eisenhart, M. (1990). Standards for qualitative and quantitative research: A prologomenon. *Educational Researcher*, 45(2), 2-9.
- Jacob, E. (1988). Clarifying qualitative research: A focus on traditions. *Educational Researcher*, 17, 16-24.

- Jacob, E. (1987). Qualitative research traditions: A review. *Review of Educational Research*, 57(1), 1-50.
- Johnson, M. C. (1977). *A review of research methods in education*. Chicago: Rand McNally.
- Kember, D., & Murphy, D. (1990). Alternative new direction for instructional design. *Educational Technology*, 30(8), 42-47.
- Kerlinger, F. R. (1973). *Foundations of behavioural research*. New York: Holt, Rhinehart and Winston.
- Kerr, S. T. (1989). There are no ends, only means. *Canadian Journal of Educational Communication*, 18(2), 143-145.
- Kerr, S. T. (1985). Asking new questions about technology and the social world of education. *Educational Communications and Technology Journal*, 33(1), 3-5.
- Knowlton, J. Q. (1964). *A conceptual scheme for the audio-visual field*. Bloomington, Indiana: Indiana University.
- Kripke, S. (1980). *Naming and necessity*. Oxford: Basil Blackwell.
- Kuhn, T. S. (1970). *The structure of scientific revolutions (2nd ed.)*. Chicago: University of Chicago Press.
- Kuhn, T. S. (1962). *The structure of scientific revolutions (1st ed.)*. Chicago: University of Chicago Press.
- Lacey, A. R. (1976). *A dictionary of philosophy*. London: Routledge & Kegan Paul.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programs. In I. Lakatos and A. Musgrave (Eds.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.
- Landa, L. N. (1983). Descriptive and prescriptive theories of learning and instruction: An analysis of their relationships and interactions. In C. M. Reigeluth (Ed.), *Instructional design theories and models: An overview of their current status* (pp. 55-68). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Laudan, L. (1977). *Progress and its problems*. University of California Press.
- Laudens, L. (1965). Grunbaum on the Duhemian argument. *Philosophy of Science*, 32, 295-299
- Lecompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1), 31-60.
- Linquist, E. F. (1956). *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin Company
- Long, J. S. (1983). *Covariance structure models: An introduction to LISREL*. Beverly Hills, CA: Sage Publications.
- Lacey, A. R. (1976). *A dictionary of philosophy*. London: Routledge & Kegan Paul Ltd.
- Losee, J. (1980). *A historical introduction to the philosophy of science*. Oxford: Oxford University Press.

- Mackenzie, B. D. (1977). *Behaviourism and the limits of scientific explanation*. London: Routledge.
- Macmillan, C. J. B., & Garrison, J. W. (1984). Using the "new philosophy of science" in criticizing current research traditions in education. *Educational Researcher*, 13, 15-21.
- Mandl, H. & Lesgold, A. (Eds.). (1988). *Learning issues for intelligent tutoring systems*. New York: Springer Verlag.
- Masterman, M. (1970). The nature of a paradigm. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.
- Mathison, S. (1988). Why triangulate? *Educational Researcher*, 17, 13-17.
- McClancy, J. (1986, May 5). The flawed tapestry. *Times Higher Education Supplement*, pp. 1, 12.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46(4), 806-834.
- McMillan, J. H. & Schumacher, S. (1989). *Research in education: A conceptual introduction*. New York: Scott, Foresman and Company.
- Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 21, 103-115.
- Merrill, M. D., Zhongmin, Li, & Jones, H. K. (1990a). Limitations of first generation instructional design. *Educational Technology*, 30(1), 7-71.
- Merrill, M. D., Zhongmin, Li, & Jones, H. K. (1990b). Second generation ID (ID<sub>2</sub>). *Educational Technology*, 30(2), 7-14.
- Merrill, M. D., Zhongmin, Li, & Jones, H. K. (1990c). The second generation instructional design research program. *Educational Technology*, 30(2), 26-31.
- Miles, M. B., & Huberman, M. A. (1984a). Drawing valid meaning from qualitative data: Toward a shared craft. *Educational Researcher*, 13, 4-11.
- Miles, M. B., & Huberman, M. A. (1984b). *Qualitative data analysis: A sourcebook of new methods*. London: Sage University Publications.
- Mitchell, P. D. (1989). The future of educational technology is past. *Canadian Journal of Educational Communication*, 18(1), 3-27.
- Mitchell, P. D. (1972). Educational Technology. In the *Encyclopedia of educational media, communications and technology*. Westport, Conn.: Greenwood Press.
- Morris, C. W. (1946). *Signs, language and behaviour*. New York: Prentice Hall.
- Neuman, Delia (1989). Naturalistic inquiry and computer-based instruction: Rationale, procedures, and potential. *Educational Technology Research and Development*, 37(3), 39-51.
- Orey, M. A., III et al. (1989). A philosophical critique of null hypothesis testing. *Journal of Research and Development in Education*, 22(3), 12-21.



- Osgood, C. E. (1953). *Methods and theory in experimental psychology*. New York: Oxford University Press.
- Peraton, H. (1987). Theories, generalizations and practices in distance education. *Open Learning*, 2(3), 3-12.
- Peraton, H. (1981). A theory for distance education. *Prospects XI*, 1, 13-24.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 347-353.
- Popper, Karl. (1969). *Conjectures and refutations*. London: Routledge & Kegan Paul.
- Polson, M. C. & Richardson, J. J. (Eds.). (1988). *Foundations of intelligent tutoring systems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Putnam, H. (1980). Models and reality. *Journal of Symbolic Logic*, 45(33), 464-482.
- Putnam, H. (1974). The refutation of conventionalism. *Nous*, 8, 25-41.
- Quine, W. V. O. (1980). Two dogmas of empiricism. In *From a logical point of view*, (pp.20-46). Cambridge, Mass.: Harvard University Press.
- Quine, W. V. O. (1963). On simple theories of a complex world. *Synthese*, 15, 103-106.
- Quinton, A. (1977). Conventionalism. In A. Bullock & O. Stallybrass (Eds.), *The fontana dictionary of modern thought*, (p. 137). London: Collins.
- Reichenbach, H. (1951). *The rise of scientific philosophy*. Los Angeles, CA: University of California Press.
- Reigeluth, C. M. (1989). Educational technology at the crossroads: New mindsets and new directions. *Educational Technology, Research and Development*, 37(1), 67-80.
- Reigeluth, C. M. (1983). Instructional Design: What it is and Why it is? In C. M. Reigeluth (Ed.), *Instructional Design Theories and Models: An Overview of their Current Status*, (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Reiser, R. A. (1988). Instructional designers in public schools and higher education: predictions for the year 2001. *Journal of Instructional Development*, 11(3), 3-6.
- Resher, N. (1979). *Cognitive Systematization*. Oxford: Basil Blackwell.
- Richards, I. A. (1923). *The meaning of meaning*. New York: Harcourt Brace.
- Richey, R. (1986). *The theoretical and conceptual bases of instructional design*. London: Kogan Page.
- Rist, R. C. (1980). Blitzkrieg Ethnography: On the transformation of a method into a movement. *Educational Researcher*, 9, 8-10.
- Rudner, R. (1961). An introduction to simplicity. *Philosophy of Science*, 28(2), 109-119.
- Schiffman, S. S. & Gansneder, B. M. (1988). Graduate programs in instructional technology: Their characteristics and involvement in public education. *Journal of Instructional Development*, 11(3), 22-28.
- Schneider, W., & Trelber, B. (1984). Classroom differences in the determination of achievement changes. *American Educational Research Journal*, 21(1), 195-211.

- Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good enough principle. *American Psychologist*, 40(1), 73-83.
- Shaver, J. P., & Norton, R. S. (1980). Randomness and replication in ten years of AMERICAN EDUCATIONAL RESEARCH JOURNAL. *Educational Researcher*, 9, 9-15.
- Slavin, R. E. (1984). *Research methods in education: a practical guide*. Englewood Cliffs, NJ: Prentice Hall.
- Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher*, 15, 4-12.
- Smith, J. K. (1985). Social reality as mind-dependent versus mind-independent and the interpretation of test validity. *Journal of Research and Development in Education*, 19(1), 1-9.
- Smith, J. K. (1984). The problem of criteria for judging interpretive inquiry. *Educational Evaluation and Policy Analysis*, 6(4), 379-392.
- Smith, J. K. (1983). Quantitative research versus qualitative research: An attempt to clarify the issue. *Educational Researcher*, 12, 6-13.
- Sneed, J. (1971). *The logical structure of mathematical physics*. Dordrecht: Reidel.
- Stewart, A. (1984). Appropriate educational technology: Does "appropriateness" have implications for the technical framework of educational technology. *Educational Communications and Technology Journal*, 33(1), 58-65.
- Suppes, P. (1957). *Introduction to logic*. Princeton: Van Nostrand.
- Sweet, R. (1986). Student dropout in distance education: An application of Tinto's model. *Distance Education*, 7(2), 201-213.
- Tennyson, R. D. (1990). A proposed cognitive paradigm of learning for educational technology. *Educational Technology*, 30(6), 16-19.
- Thomas, D. (1979). *Naturalism and Social Science: A Post Empiricist Philosophy of Science*. Cambridge: Cambridge University Press.
- Tight, M. (1988). Defining distance education. *International Council for Distance Education Bulletin*, 18, 56-60.
- Tinto, V. (1982). Limits of theory and practice in student attrition. *Journal of Higher Education*, 53(6), 687-700.
- Tranel, D. D. (1981). A lesson from the physicists. *Personnel and Guidance Journal*, 59(7), 425-428.
- Travers, R. (1964). The transmission of information to human receivers. *AV Communications Review*, 12, 373-385.
- Tuckman, B. M. (1978). *Conducting educational research*. New York: Harcourt, Brace, Jovanovich.
- Van Frassen, B. (1980). *The scientific image*. Oxford: The Clarendon Press.
- von Bertalanffy, L. (1968). *General systems theory*. New York: George Braziller.

- von Neumann, J. (1955). *Mathematical foundations of Quantum Mechanics*. Princeton: Princeton University Press.
- Wedeking, G. (1969). Duhem, Quine and Grunbaum on falsification. *Philosophy of science*, 36, 375-399.
- Weldon, T. D. (1957). *The vocabulary of politics*. Harmondsworth, Middlesex: Penguin.
- Wenger, E. (1987). *Artificial intelligence and tutoring systems: Computational and cognitive approaches to the communication of knowledge*. Los Altos, CA: Morgan Rantmann Publishers, Inc.
- White-Beck, L. (1975). *The actor and the spectator*. Boston: Yale University Press.
- Winn, W. (1989, October). *Is there a theoretical basis for Educational Technology?* A public lecture presented at Concordia University and sponsored by the Graduate Programme in Educational Technology, Montreal, Quebec.
- Wolcott, H. (1980). How to look like an anthropologist without really being one. *Practicing Anthropology*, 3(1), 6-7, 56-60.
- Woodger, J. H. (1957). *The axiomatic method in biology*. Cambridge: Cambridge University Press.
- Woolfolk, A. E. (1987). *Educational psychology (3rd. ed.)*. Englewood Cliffs, NJ: Prentice-Hall.