CONCEPTUAL THERAPY:

AN INTRODUCTION TO FRAMEWORK-RELATIVE EPISTEMOLOGY

STEVEN J. BARTLETT

CONCEPTUAL THERAPY:

AN INTRODUCTION TO FRAMEWORK-RELATIVE EPISTEMOLOGY

STEVEN J. BARTLETT

Website: http://www.willamette.edu/~sbartlet

STUDIES IN THEORY AND BEHAVIOR
Saint Louis, Missouri

Printed in the United States of America
First printing, 1983

eBook edition, 2014

To err is human;
To remain in error is crazy.

(Found in a Chinese fortune-cookie)

Human history presents a spectacle of
the repeated failure of great ideas
to penetrate the human heart.

(Jacob Needleman: The Heart of Philosophy)

Note to the reader

Blank pages contained in the first edition have been allowed to remain in this eBook edition in order to preserve the original page numbering.

CONTENTS

Acknowledgements

The belief that the subject-matter of this book is relevant, important, and interesting is due in large measure to the encouragement I have received over the years from three men.

Robert M. Hutchins, whose presence in the world is now sorely missed, offered an example of dedication through his life and his work, to the ideals of what he liked to call The Great Conversation, the dialogue of the past two thousand years which is the foundation of western civilization.

I am indebted to Paul Ricoeur, who gave his strong support to the early thought this book rests on. He is a true scholar, a generous teacher, and a person of rare integrity.

I would like to express my thanks to C. F. von Weizsäcker, former Director of the Max-Planck-Institut in Starnberg, West Germany, for his invitation to continue my research in Starnberg during the period 1974-75.

My indebtedness of course goes beyond these men.

I am grateful to my colleague and good friend, Raphael J. Becvar, whose general systems approach to family therapy greatly encouraged my interest in building bridges between epistemology and therapy.

Finally, and perhaps most of all, I am indebted to my mother, Elizabeth Bartlett, whose sharp intellectual curiosity and excitement in the company of great ideas were conveyed to me in many formative philosophical conversations which usually lasted hours beyond midnight.

Acknowledgement is gratefully made of the following journals, from which these papers by the author are reprinted:

"Towards a Unified Concept of Reality", ETC.: A Review of General Semantics, Vol. 32, No. 1, 1975, pp. 43-49.

"The Idea of a Metalogic of Reference", Methodology and Science, Vol. 9, No. 3, 1976, pp. 85-92.

"Phenomenology of the Implicit", Dialectica: Revue internationale de philosophie de la connaissance, Vol. 29, Nos. 2-3, 1975, pp. 173-188.

"Self-Reference, Phenomenology, and Philosophy of Science", Methodology and Science, Vol. 13, No. 3, 1980, pp. 143-167.

"Referential Consistency as a Criterion of Meaning", Synthèse, Vol. 52, 1982, pp. 267-282.

Background

The main ideas in this work have been a long time incubating. Many can be traced back twenty years, when I was majoring in physics as an undergraduate. At that time, it seemed to me that many of the basic principles we were taught to use in solving physics problems were doubtful and frequently unclear, through no fault of my teachers. It gradually became evident to me that the questions I wished to ask could not be answered in physics, for they were questions about the foundations of physics, which physicists must presuppose in order to get on with their tasks. In my last year in college, I therefore decided to shift my major to philosophy, hoping to find there a metaframework which would make possible the reflective enterprise I wished to participate in.

Unfortunately, I soon discovered that philosophy is made up of competing philosophies, in any one of which the practitioners of that branch seemed reluctant to ask foundational questions about their own approach. Phenomenology, at least the phenomenology of the young Husserl, captured my interest and taxed my patience for about seven years, seeing me through graduate school and my first university teaching appointment. Phenomenology promised to provide a meta-framework of the sort I wished for, one that could be used to examine critically and reflectively the special theories of the sciences, and, at the same time, provide a clarity of self-understanding that would make the choice of the phenomenological perspective non-arbitrary, and, in some sense, self-justifying.

However, after seven years of effort, it seemed clear to me that this promise of Husserl's phenomenology was still a promise -- in fact, a group of promises of a metaframework that had still to be evolved. During the more than half a century since Husserl published the first volume of his Ideen in 1913, phenomenology had fragmented into many provincialisms, some existentialist, some more methodologically oriented, but none any closer, I felt, to fulfilling the original set of promises. Too, the terminology of phenomenology, which began badly in Germanic convolutedness, ended badly: it was excessively inbred, top-heavy, and unnecessarily obscure. Phenomenology had itself become a branch of philosophy, one of the competing alternative paradigms, rather than a neutral metasystem for critical investigation. And so, disappointed, I decided to try to make my own path.

My dissertation was a first attempt. It sought to use some of my phenomenological studies as a basis, in order to develop a self-referential metatheory of the sort I was aiming at. Since I found the vocabulary of concepts in phenomenology problematic in its own right, I set out to construct my own. It was, in retrospect, something of a philosophical private language: at least the

concepts I tried to develop and name were familiar only to a minority of one! It was a tribute to French open-mindedness, the true spirit of the liberal arts, and in particular a tribute to Paul Ricoeur, involved in creative work, to tolerate and even encourage efforts that claimed to stand against so much of tradition and against much that defined philosophy. This was more than a decade ago.

Since then, I have tried to translate this early, technical vocabulary into a less specialized, more open system of concepts. The result is something I have come to call "conceptual therapy". Whether it can serve as a basis for a neutral metaframework for philosophy as well as other disciplines, the tests of time will determine.

Foreword

The history of science and technology could be re-written to show that their source is man's intellectual and physical laziness. As Whitehead pointed out, the whole object of science is to avoid thinking as much as possible.

The progress of science coincides with progress in the effectiveness of its models. A model is a simplified representation of (usually only some aspect of) reality. Science concerns itself with investigating its own models, their predictive capabilities, their descriptive and explanatory power, their consistency. When a model shows signs of misfit, when the data do not conform to the expectations of a theory, the model is revised or discarded.

Already, this is an admission that we require simplified representations to cope with a complex reality. It is also an admission that we expend our efforts in working on those representations, improving and developing them. We don't confront reality "head on".

Our models and the mathematical tools we use to develop, analyze, and evaluate them, are all in the category of intellectual crutches. They serve us precisely because the thought that has been invested in them reduces the amount of subsequent thought required to organize information.

All of the major achievements of mathematics, natural science, and technology can be seen in this light: their purpose can be construed as attempting to reduce effort. Information has now been organized under such categories as mathematical tools, scientific laws, engineering techniques, designs, etc. Their usefulness lies in the fact that they do not need to be reinvented every time they are needed.

As a species, we seek compulsively to avoid thinking and other activity that expends effort. We will work inordinately hard to reduce the work involved in subsequent work!

In spite of our lazy nature, we get along very well without thinking a great deal. We memorize formulas, laws, principles, facts -- precisely in order to get by with thinking less. That we manage to get along in our environment as well as we do, is a tribute to our own powers of reducing the difficulty of tasks. It is all motivated by a distaste, even a genuine abhorrence, for work, in particular mental work!

This is important, for I do not believe we are very good at thinking. There is a lot of evidence for this, of both the experimental and the everyday sort. Stupidities abound and proliferate. Usually they are someone else's, of course, but they no less commonly can be our own.

Epistemology is a discipline that requires much conceptual effort. It involves thinking about thinking, thinking about what we can and cannot say, correctly, that we know. Epistemology has not, at least as yet, been mechanized in any sense: There are no known "effective decision tests" that will tell us if a certain claim to knowledge is epistemologically well-founded. Although generalizations can sometimes be made, these are of little value unless compelling epistemological analyses can be developed, usually pretty much from scratch, pertaining to the case at hand. This often demands some originality, to produce a "custom fit".

Most approaches to epistemology describe specific epistemological analyses and their results. They are information-oriented. They ask the reader to read, and to memorize. This is exactly what thought-avoidance would prescribe.

The approach described here is the opposite in perspective. It concentrates on examining a group of methods that the reader can learn to apply himself in whatever area he may be interested in: science, psychology, mathematics, literary criticism, anthropology, etc. The approach here is skill-oriented, with an interest in the practice of epistemology, rather than its history.

To the Teacher

This work is an expression of the conviction that epistemology, in an extended sense of the term, should be practiced by every thinking being. Its applicability goes considerably beyond philosophy. Its proper place, according to the perspective presented in this book, is somewhere between the development of cognitive skills and conceptually-oriented therapy.

As a result, my approach in this work is non-standard. There is no attempt to introduce students to epistemology by means of the history of philosophy. What is drawn from the history of philosophy will be evident mainly in the choice of examples, and in passing references to Kant, to a few members of the Vienna Circle, and to early phenomenology. History is not the framework this book is about; rather, it is about the practice of epistemology.

For some, it may cause eyebrows to raise, to consider teaching practical skills in epistemological analysis to undergraduate students, many of whom have had little and sometimes no previous exposure to philosophy. I attest to the fact that this is what I have done with a good measure of success, from what I can tell, in various of my undergraduate as well as graduate courses for more than a decade. I have taught epistemology, in one form or another, in the context of classes in philosophy of physics, mathematics, and behavioral science, in classes in positivism, phenomenology, and continental thought, and, of course, in classes in epistemology. I have frequently found non-philosophy majors to be highly motivated, perhaps because there has been little "hardening of the philosophical categories" in their case, leaving them with fewer axes to grind, or less devotion to a particular creed to cloud their vision. No matter what their major, many students have welcomed the opportunity to step back from their habitual frames of reference and become involved in the critical examination which epistemology encourages. I have come to believe, as a result, that epistemology has an interested and wider audience than it traditionally has been thought to have.

The effort in these pages was motivated by a desire to describe a pluralistic yet critical, comparatively neutral framework in terms of which practical skills in epistemological analysis can be studied and developed. Emphasis is placed on self-referential argumentation, which I believe to constitute one of the most powerful approaches to conceptual evaluation.

The book takes seriously the notion that many of our ideas and theories stand in need of therapy, because they frequently are in conflict with what we intend to do. We are, as a result, often trapped in self-undermining projects.

The pages that follow were written for my students. I have tried to introduce the subject in a way that does not lose sight of the forest for the scholar's trees. I have refrained from normal academic style with its many qualifications and mention of alternative positions. The account of conceptual therapy, its purposes, techniques, and practice, is told directly.

The approach that is described can be used in the context of a wide range of courses which encourage students to develop their own reflective and analytical abilities. There is no good reason why conceptual therapy should be restricted to philosophy. Other fields -- psychology, education, research methods, counseling

and psychotherapy, the natural and social sciences, etc. -- can gain much by epistemological self-consciousness. Already, it appears as if epistemology may come to form a field of study independent of philosophy from which it branched.

Occasional exercizes are interspersed in the text. These are few in number and are not intended to make up a sufficient set of exercizes in epistemological analysis. The object is to introduce students to a group of techniques which can then be applied to the specific framework of interest in their class, whether it is philosophy of physics, mathematics, behavioral science, another branch of philosophy, or another discipline entirely. These latter specialized applications will give students the practice they need to develop skills in epistemological analysis.

I have used earlier drafts of this book, and portions of it, in various courses, as already mentioned. In undergraduate courses, I use Part I only; in graduate seminars, students read both parts. In courses in the philosophy of science, for example, I usually teach elementary proof-techniques of symbolic logic, discuss the development of criteria of meaning, and then go on to deal with epistemological techniques of analysis. Together, these occupy perhaps a third of a semester. The remainder of the semester is devoted to applications of these techniques to arguments and positions identified by members of the class, drawn from expository works in physical science, mathematics, or behavioral science, or the philosophies thereof. In my courses in philosophy of physics, e.g., I ask students to apply epistemological techniques of analysis to works which examine the nature of physical theory, the principle of minimum assumption, theory construction and models, theories of measurement, the role of the observer, probability and induction, theory of explanation, causality, relativity, quantum theory, etc. In courses in philosophy of behavioral science, epistemological analyses are developed which relate to the covering law position, reflexive predictions, behavioral reductionism, general systems and behavior modeling, learning theory, analyses of various approches to psychotherapy, etc. In courses in philosophy of mathematics, epistemological analyses are made of specific formalist, logicist, and intuitionist positions, of the psychology of mathematical invention, of various of the limitative theorems, etc.

Whether at the undergraduate or graduate level, it is comparatively easy to bring attention to bear on basic claims to knowledge in a particular area of study. In seminars in phenomenology, fruitful applications of epistemological techniques of analysis can be made concerning the phenomenological approach itself and the specifically phenomenological means available for justifying its knowledge claims. Similar reflective and evaluative studies can be made of the positivists' criteria of meaning, of the theses propounded by British analysts or Continental philosophers, etc. The evaluative framework of epistemology is a general one, which can be employed in assessing the justifiability of virtually any of man's claims to know, in whatever disciplines these may be formulated.

In my experience with students who often have had little or no previous relevant experience, many have been able to produce systematic, carefully developed epistemological analyses of good quality. The skills they learn, perhaps they will retain somewhat longer than factual memory endures, when, "after they have forgotten everything they used to know", they are left with the fruits of their education.

PART I

THE THEORY AND PRACTICE OF CONCEPTUAL THERAPY

Our imprisonment consists exclusively of the fact that we are not aware of being in our prison. So we cannot (logically) both be in this condition and know that we are in it, and knowledge of our condition is instantly delivering, like the cure for a disease which consists only in not having the cure. But deliverance is complicated by the extraordinary difficulty of explaining to the prisoners, in terms intelligible to them, that prisoners are what they are. For the conditions which make self-understanding possible are incompatible with the conditions they are in, and he who speaks of imprisonment to prisoners must be regarded by them as a madman in his raving. For the bonds which hold them captive are the boundaries of the understanding, and how are we to bring the boundaries within themselves to make them understood? The limits of understanding are not part of what is understood.

(Arthur C. Danto: What Philosophy Is)

CHAPTER 1

Conceptual Therapy

Introductory Note

The notion that we at times need therapy, whether emotional, physical, or occupational, is commonly accepted. The analogous notion that in our use of concepts we at times also are in need of therapy, is less widely accepted, though still familiar. Plato's Cave, Wittgenstein's fly-bottles, and Ryle's category mistakes come to mind. Toulmin has coined the word 'cerebroses' to refer to conceptual neuroses. John Wisdom calls Wittgenstein's later philosophy a kind of intellectual psychoanalysis. Bateson speaks of pathologies of epistemology.

In psychotherapy, it is often appropriate to regard as pathological a person's rigid commitments to patterns of inconsistency. Similarly, philosophy, undertaken as conceptual therapy, can serve to identify and treat (i) concepts which are incompatible with their presuppositional bases, and (ii) ways of using concepts which are self-refuting.

The materials collected together here explore a pair of analogies:

(a) As human behavior at times becomes self-defeating and in need of psychotherapy, so sometimes do human concepts stand in need of therapy.

(b) In somewhat the sense in which theories of psychotherapy express forms of therapy that are used to treat self-defeating behaviors, so is a general epistemological therapy for dysfunctional concepts possible.

Although conceptual therapy and theories of psychotherapy constitute independent areas of interest, there are firm connections between them. While it is true that conceptual therapy is conceptual and impersonal in nature, its ultimate application is the conceptual structures of individuals. Both psychotherapy and conceptual therapy seek to improve man.

Conceptual therapy can be of some use to psychotherapy and psychotherapists, since theories of psychotherapy and their representation in the minds of individual therapists are themselves conceptual structures. Often these structures, human in origin as they are, stand in need of therapy.

A general therapy for concepts can respond to a concern among some philosophers and psychotherapists to develop ways to study the theoretical foundations of psychotherapy. Conceptual therapy seeks to provide a metaframework that can do this for any discipline.

Overview

Man's evolution can be measured in terms of the growth of his knowledge. "Man, by his nature, desires to know" asserts Aristotle in the opening sentence of the Metaphysics. If man desires to know, he is inclined even more to wish to claim that he knows more than he does. It was in connection with his unacknowledged lack of modesty that the Socratic mission was born: The epistemologist, historically, is foremost a "gadfly" who, with a singular lack of charity, proposes to sort claims to knowledge into two categories: on the one hand, one can and does know what one is talking about; on the other, one cannot and does not know what one is talking about.

Thomas Harris once said: "to say that we are free is merely to mean that we know what we are doing." A host of tragedies and errors is the inheritance of people of whom it is truly said that "they know not what they do."

To provide us with the freedom which accurate self-knowledge brings with it, epistemology studies three interrelated subjects: (i) the limits of knowledge, (ii) the (necessary and sufficient) conditions of knowledge, and (iii) the preconditions or transcendental presuppositions of knowledge.

When we know what the limits of knowledge are we are able to sort knowledge claims into two categories: specious, or empty, and genuine. We are able to justify our decisionswhen we classify claims in this way, and hence avoid introducing judgements which reflect no more than personal opinion.

When we know the conditions of knowledge, we have answered the question, "Under what circumstances are we justified in claiming that we know?" We are able to determine, for a given knowledge claim, whether it satisfies certain of these conditions.

When we know the preconditions of knowledge, we understand what must be presupposed in order for knowledge to be possible at all. If we do not take these preconditions into account, knowledge ceases to be possible. Transcendental presuppositions are essentially unavoidable if we are to be able to know.

Detecting the limits of knowledge is a critical, evaluative task. Identifying the conditions of knowledge is a descriptive, analytical undertaking. Showing that certain presuppositions are unavoidable for knowledge of a particular variety to be possible is a demonstrative, argumentative endeavor.

This work contains a group of materials which seek to clarify some of the means available to us if we wish to undertake these tasks. Specifically, it contains:

- A pair of psychological tests which should help you to become aware of the kind of reflective, critical, careful thinking epistemology requires.

- A description of epistemology as a discipline which seeks to remove certain kinds of "conceptual blocks" which imprison thought in self-sabotaging patterns.

- A summary of some of the main tools, models, or approaches which the history of philosophy provides us and which are useful in justifying our claims to know (i) the limits of a knowledge claim, (ii) its conditions, and (iii) its pre-conditions. Without these or similar tools, our claims to know these things would degenerate into statements of opinion or personal judgment.

- A description of a format for a post-Kantian approach to transcendental deduction. The objective here is to make clear how one can prove that a given knowledge claim indeed entails certain preconditions, without which what is expressed could not be expressed and without which what is claimed to be known could not be known.

- A review of basic principles of self-referential argumentation.

- A group of axioms for framework-relative epistemology. It is argued that one cannot <u>not</u> accept these basic statements, on pain of incoherence, and that hence they provide us with a useful and certain starting point for epistemological analysis.

- A view of the relationship between the approach to epistemology described, and the concept of the double-bind.

- A set of sample epistemological analyses.

- A number of exercizes and problems.

- A summary of some of the most common epistemologically problematic claims and concepts you are likely to encounter whether in philosophy, or in non-philosophical theories about specialized subject-matters, e.g., behavioral science, mathematics, physics, etc.

- Some guidelines that may be useful to you in constructing assessments of arguments.

- A group of guidelines for epistemological analysis.

- A description of certain main ways in which knowledge claims <u>can</u> be justified.

- A group of five papers which contain more detailed studies and illustrations of the self-referential approach involved in conceptual therapy.

The Idea of Conceptual Therapy

Everyone who is raised and lives within a culture develops patterns of intellectual response which take their meaning from the general environment of ideas of the culture. These patterns of intellectual response reflect a person's belief system; it makes up the basic set of beliefs which a person's thoughts and behavior are built upon.

A person's belief system is brought most strongly to awareness when certain of his philosophical or ideological commitments are called into question. The cultural background establishes a general vocabulary of ideas; an individual's belief system is constructed from a sub-set of these; and his philosophical or ideological commitments express particular claims which he advocates, usually unquestioningly, based on his system of beliefs.

These intellectual commitments are usually very rigid, largely because they have been accepted without question, often as an implicit ingredient in one's upbringing and education. They are not easily opened to reflective questioning, and are less easily changed even when they do not serve the individual's intended purposes. Some commitments are literally self-defeating: They block the attainment of an individual's personal objectives and interests. (A short-circuit is a concrete example of a self-defeating system.) Theories, too, as we shall see, can be founded upon self-defeating commitments.

When an individual's unexamined commitments are self-defeating, he will be inclined to accept, recommend, and base decisions on claims which are self-undermining. His commitments, as well as the moral and practical inclinations to which his commitments lead, will foster self-defeating attitudes and behavior.

Conceptual therapy attempts to deal with self-undermining claims and concepts, whether they are encountered in the belief systems of individuals, or in the conceptual structure of theories.

An individual who requests psychotherapy usually does so because he is having difficulty with others, his job, etc., and is in pain in various ways. If his patterns of behavior have become a rigidly defined tangle of dissatisfaction and frustration, and he cannot free himself to establish alternative ways of behaving, we say his trap is psychopathological. No matter how dysfunctional his patterns of behavior may have become, yet they are his, and they are the only ones he knows. So that, even when he seeks out therapy, there is usually a point that is reached in therapy when he will resist change.

In conceptual therapy, on the other hand, the intellectual commitments that are called into question usually have not been considered to be a possible source of discomfort or dissatisfaction. They are not known to be dysfunctional at all, even when they are. The more rigidly a person is committed to his belief

system, the more stubbornly he will resist questioning it.

To persuade a person, or the advocate of a particular theory, to question his belief system, various conceptual puzzles are used to ensnare him, for which philosophy is well-known. But unfortunately, even if he should come to find philosophical investigation intrinsically interesting and rewarding, the individual all too easily acquires one philosophical belief system of any of various schools of thought, to which he now adheres as rigidly as he did to his old, unquestioningly held set of beliefs.

The effectiveness of conceptual therapy becomes possible when the very beliefs an individual subscribes to can be used to reveal to him their self-defeating nature. It is effective precisely because no external standards of normality or justification are relied upon.

In psychotherapy, the usual focus is on a disturbed individual or family group. In conceptual therapy, on the other hand, the focus is on an impersonal representation of some particular set of conceptual commitments. A large number of individuals who share that conceptual structure may therefore come to be addressed implicitly.

Conceptual therapy accepts the view that an entire culture may be based on beliefs many of which are epistemologically pathological. Conceptual therapy suggests that many of man's cruelties, his capacity for viciousness and hatred, stem from such self-defeating conceptual commitments. For example, the dogmas of numerous groups, each of which claims to have found the truth, independently of all conceptual frameworks, set men in ideological opposition to one another. Unfortunately, contention between exclusivist ideologies frequently goes beyond the verbal, and results in persecution, destruction of lives, and great suffering and unhappiness. Warfare is the open expression of an intellectual and moral short-circuit in communication between nations: Paths of communication are blocked by mutually exclusionary, equally self-righteous dogmas that are accepted, usually blindly. War is a mental illness that affects an entire population. It is man's most self-destructive behavior, and it stems, always, from a breakdown of communication between contending and rigid ideologies.

In conceptual therapy, the views expressed by an individual or group receive explicit attention. Most people construct or organize their sense of self-identity around their conceptual commitments; the belief-system to which they daily swear tacit allegiance comes to constitute much of their sense of personal identity. To this extent, reflective criticism which questions the impersonal representation of a set of conceptual commitments to which an individual rigidly adheres, cannot help but be perceived by him as a threat to his personality structure, as an attack on his intellectual identity, or his moral sense, or his practical investments and personal involvements. It is an unusual person who possesses the creative capacity to embrace the results of reflective self-evaluation and grow beyond the warm, blind comfort of unexamined dogma. Although the focus of conceptual therapy is on an impersonal representation of a system of intellectual commitment, there is an unavoidable

ad hominem dimension which may be experienced as intimidating. If change is to be effected, both psychotherapy and conceptual therapy require trust, confidence in one's therapist, which is fundamentally a confidence in one's own judgment. An atmosphere is essential which fosters creative development without compromising the standards of dispassionate reflection.

Albert Ellis is a psychotherapist who developed a theory of personality called rational-emotive therapy during the 1950's. Ellis claimed that people tend to create their own emotional reactions. He observed that they tend, as time goes on, to exacerbate, rather than lessen, this tendency. However, they do have the ability to understand what is disturbing them, since they have a distinctive capacity to think about their own thinking. They can train themselves to change or eliminate their self-sabotaging beliefs, because they have a capacity for self-discipline. If they work hard at understanding their previously unexamined belief systems, they can exercize considerable control over their disturbance-creating tendencies. If they are helped to become aware of their "crooked" thinking and behavior by a philosophic, highly directive, didactic therapist (who may or may not be a warm empathizer), they are much more likely to change their symptom-creating beliefs.

Rational emotive therapy is foremost a theory of personality change. It is to psychotherapy what conceptual therapy is to epistemology. Ellis found that human beings have exceptionally powerful built-in tendencies to think irrationally and to harm themselves and others as a consequence. Among their tendencies to be self-sabotaging, Ellis identified the following factors; in the righthand column are analogous characteristics which concern conceptual therapy:

Characteristics inclining people toward self-sabotaging behavior and attitudes:	Parallel conceptual characteristics:
1. Inertia, laziness, resisting change because of the effort needed to effect it.	1. Mental inertia, laziness; reflection is hard work to be avoided.
2. "What I desire I must really need."	2. "If my unexamined biases are comfortable and I like them, they are true."
3. Rut-making: inability or unwillingness to initiate alternative patterns.	3. Inability or unwillingness to reflect on one's premisses, to step outside of one's current frame of reference.
4. Rigid adherence to early training: religious preferences of parents, political ideas, etc.	4. Rigid adherence to "received" conceptual system.
5. Overcaution; defensive anxiety when made aware of alternative viewpoints.	5. Resisting new ideas, manifestly conservative.

6. "Need" (i.e., desire) to be superior.	6. Framework-centricism: provincialism, conceptual megalomania or absolutism, "I have the truth."
7. Extremist oscillations in thinking.	7. Confined to thought involving polar opposites: If idealism is wrong, then realism has got to be right.
8. Reliance on automatic habits of thought and behavior.	8. Reliance upon familiar concepts and habitual ways of thinking.
9. Tendency to forget easily what is harmful.	9. Continued use of a concept or framework even when it is evidenced as self-destructive.
10. Commitment to wishful thinking.	10. Commitment to the independent reality of objects which are understandable only relative to a specific frame of reference.
11. Inability to achieve self-discipline, tendency to defer, to procrastinate.	11. Inability to transfer conviction from theoretical proofs to practical decisions.
12. Expects just treatment, dwells on disappointments.	12. Expects all ideas and experience will reinforce biases; is continually upset when this is not the case.
13. Total self-condemnation when only some aspect of behavior is erroneous.	13. Readiness to give up reflection if it shows that his belief system is partially incoherent.
14. Tendency to overgeneralize, to exaggerate the applicability of a single approach.	14. Tendency to overestimate the range of applicability of habitual concepts.
15. Adherence to a rigid belief system encourages emotional disturbances in some situations, and this often is followed by physical disturbances, i.e., psychosomatic symptoms.	15. To convert reflective criticism of belief system into a less painful form, tendency is toward a schizophrenic reaction: to engage in fantasizing and so make the reflective criticism unreal (hebephrenia); to shut out the criticism and assume the ostrich stance (catatonia); or to construe criticism as hostile, and react with fear or anger (paranoia).

It is evident that this list emphasizes undesirable characteristics, which block positive personal or conceptual change. Unfortunately, most of us, in

one respect or several, have these characteristics. If we were to contrast the above description of a self-sabotaging person with a more positive description of a self-actualizing individual, we would include such characteristics as:

- confidence in one's own decision-making and reflective abilities
- an interest in intellectual, moral, and personal development
- a suspicion of dogmas prescribed and backed by authority and tradition, because they so easily influence
- a comparative freedom from fear of the unknown, from fear of being an individual, from fear of non-conformity
- a capacity to adapt easily to new experience, i.e., a high degree of intellectual and personal flexibility
- an interest in placing one's habitual frame of reference at a distance, in order to evaluate its strengths and weaknesses
- a tolerance for self-criticism, hence a kind of moral courage
- the enjoyment of challenge
- respect for the integrity of alternative approaches; a tolerance for pluralism; dissatisfaction with provincialisms
- a positive interest in reducing self-destructive tendencies in oneself and others
- a good sense of humor.

It is likely that if this list of characteristics were taken out of its context here, it would, by many people, be seen as a definition of "creativity".

To summarize: We have compared some of the similarities between psychotherapy, particularly one theory of personality change, rational-emotive therapy, and an approach to epistemological change, conceptual therapy. Both acknowledge that human beings are suggestible, greatly influenced by their cultural and intellectual environments, and posses an incredibly facile and stubborn propensity to fail to use or to misuse reason and to invent constructs and subsequently believe in their reality. These tendencies frequently sabotage their own physical, mental, or moral health and happiness, and can be destructive to the lives of others around them. The self-destructive patterns that come to be established, whether on the psychological level or on the theoretical level, are rigidly adhered to, and are often changed or given up only after much resistance. Both rational-emotive therapy and conceptual therapy, although having distinct and divergent foci, ultimately confront the individual person with the reasons why many attitudes, behavior, or conceptual commitments

are in need of change or replacement. At this point, both therapies are effective, or fail, depending upon the willingness of the individual to accept the momentary risks which change may threaten. Ultimately, both forms of therapy presuppose that many and perhaps most individuals have a remarkable capacity to rise to challenges which force them, sometimes, to engage and transcend patterns which have become habitual, automatic, and which they have usually never questioned before.

CHAPTER 2

Detecting Conceptual Short-Circuits:

Preliminary Self-Tests

Detecting Conceptual Short-Circuits:

Preliminary Self-Tests

The following self-tests may or may not be "statistically reliable diagnostic instruments" in the context of conceptual therapy. Psychometrics has become a specialized field unto itself, into which only a few fools and philosophers venture without professional certification. The pair of self-evaluations which are included in this chapter are here not because I believe they can indicate the presence, or absence, of some skills useful in detecting and eliminating epistemological pathologies. (I do, incidentally, believe this.) They are included here rather because they give the reader an experience of the approximate sort of mental set, or analytical disposition, which is basic to the approach that is later described.

The reader is therefore invited to try the following two "tests" in a casual spirit of intellectual game-playing curiosity, for they are here and there tongue-in-cheek, and here and there not.

Directions for self-test #1:

(a) Find "mistakes" or "errors" in the following story.

(b) Explain why you believe they are "mistakes", and how you would correct them.

(c) Try to characterize the mistakes by kind, grouping cases of similar kinds of mistakes together under one class heading.

(d) Do not read beyond the end of the story until you have had a few minutes to try your hand at (a) - (c) above.

A Nice March Day

On a nice March day our club organized a day's outing. Although it had been raining all night, the roads were quite wet and muddy in the morning; but this did not spoil our pleasure. We came through a wood consisting entirely of fir trees. Unfortunately, there were no leaves on the trees as it was still early in the season; how lovely this wood must be in summer time when the trees with their shade protect us against the sun. Then we saw in the distance a rabbit. I ran after it, but as it ran faster than I did, I could overtake it only slowly and finally caught it. I did not hurt it, but let it go again shortly.

Then we came past fields where the farmers were harvesting. At noon we arrived at the village where we intended to stay. About one year ago the village had suffered heavily from a fire; the tower of the church had burned down entirely; in memory of this event, there was a tablet at the place where the top of the tower had been. In a dairy we asked whether we could have milk and cheese, but as they didn't have any at that time they told us they would milk the cows in half an hour and that then we could have plenty of both. Of course, we were glad to wait and then enjoyed it very much. Most of the crowd now remained in the village, but together with a friend, I made a little jaunt up to a vantage point. We climbed up for half an hour, enjoyed the lovely view and then returned to the village on a path which was even steeper and also continually uphill.

With many games the afternoon passed quickly. We hardly noticed that the shadows grew shorter and shorter and we were surprised when we saw that the sun was setting. We sat down for a while on the shores of a lake. All of a sudden, a dense fog came up from the lake, but it did not spread over a large area; it just covered us and all the things close to us disappeared but all the objects in the distance remained distinctly visible. Tired, but well satisfied, we reached home after complete darkness had fallen.

***

Now that you have completed "A Nice March Day", perhaps you observed yourself going through several phases, as your mind-set developed and took shape. Initially, you were probably confused about what sorts of "errors" or "mistakes" there might be. As you read along, certain inconsistencies probably jarred you out of the story-line; these little shocks of recognition (which literally can mean a kind of intellectual "double-take", as in 're-cognition') may have been accompanied by a growing feeling of amusement. -- Much humor takes advantage of an analogous developing sense of tension which is finally broken by a punch-line that is somehow inconsistent with one's expectations.

The perception of inconsistencies is one of the basic skills involved in conceptual therapy. An epistemologist does not just sit back and enjoy the "story"; instead, he is constantly evaluating the compatibility between what is actually said, what was said, what is and has been intended, what is presupposed, etc. Most of us are usually pretty much one-dimensional in our perceptions of meaning -- unless we have been endowed with wit, or read a lot of poetry, or are at least a little crazy!

Punsters are universally held in rather low repute, because their quick repartees not only show an enviable agility with meanings and word-sounds, but because they can be so unsettling! A pun produces momentary confusion among one-dimensional thinkers, and so we are momentarily and literally "put out" -- put out of the system of meanings we had intended to remain firmly embedded in.

Poetry, the divine madness, manipulates dimensions of meaning through symbolism, imagery, analogy, allegory, word-sounds, meter, etc. Forms of madness not so divine, but closer to home, and often very painful, include the paradoxical dysfunction, schizophrenia. A schizophrenic may have an uncontrollably creative capacity to perceive dimensions of unintended meaning that can lead to terrible confusions, a sense of disorientation and distortion which is difficult, and sometimes impossible, for therapists to break into in order to help a patient regain a stable perspective. Schizophrenics may suffer from an impairment of meaning-perception, so that everyday behaviors, gestures, greetings, etc., can no longer be interpreted correctly.

People who are trapped in one-dimensional universes -- who frequently have little or no sense of humor, who are often absolutists, dogmatic, rigid -- find epistemological analysis difficult. Schizophrenics and others whose reflective powers of meaning-discrimination are impaired or are uncontrollable cannot easily detect what most of us perceive as blatant incongruities in the "Nice March Day" story. It is therefore perhaps understandable why the story was originally developed and used as a diagnostic instrument by J. McV. Hunt in 1935, for just this purpose: to identify the victims of pathological entrapment. The patients he was interested in diagnosing were caught in rigid channels of thought: certain paretics (patients suffering

from certain forms of paralysis) and schizophrenics. (See American Journal of Psychology, 47, pp. 458-463: "Psychological Loss in Paretics and Schizophrenics".)

Directions for self-test #2:

Not all of an individual's belief system relates to specifically philosophical or epistemological issues and claims. Much in a belief system is tied up with personal values, expectations that have arisen in the context of interactions with family members, and other basically psychological-emotive concerns. But most people also adhere to a set of beliefs which they do not normally have occasion to think about explicitly and which provide much of the subject-matter of epistemology.

The self-evaluation that follows is, like the first "test", included here to point the reader in a certain direction, to encourage him to reflect about his own tacit philosophical beliefs, some of which relate to his sense of self-identity, his perspective on other people, his understanding of material objects, etc.

A list of claims appears below, grouped together beneath a number of traditional philosophical headings. For each claim, decide whether you would tend to agree or disagree. An answer sheet is provided, to encourage you to make a record of your present philosophical belief system. Later, as you learn techniques of epistemological analysis, it may be interesting to you to refer back to this record, in order to evaluate your original position in a new light.

Analysis of Philosophical Belief System

In relation to each of the claims that follow, indicate your agreement (Yes) or your disagreement (No). If you disagree with one or more stated or implied assertions within a single claim, you should indicate disagreement with the entire claim. If you do not disagree with any stated or implied assertion in a claim, you should indicate agreement with the claim. If a claim seems ambiguous to you, the interpretation of the claim is up to you.

1. a.
   b.
   c.

2. a.
   b.

3. a.
   b.
   c.

4. a.

5. a.
   b.
   c.

6. c.
   b.

7. a.
   b.
   c.

8. a.
   b.
   c.

9. a.
   b.

10. a.
    b.

11. a.
    b.

12. a.

13. a.

14. a.
    b.

++

1. Self

   a. There is an ego, an "I", a "self" which stands apart from the individual contents of experience, essentially distinct from individual perceptions, sensations, memories, and other psychological states.

   b. At least some of my beliefs, perceptions, sensations, or behaviors are determined by "unconscious" processes or factors over which I have no direct control and no direct, conscious access.

   c. The structure of my perceptual system, the principles of its physiological organization, my sense of time, and other structurally fundamental aspects of my consciousness act as a "sieve" or "grid" which filters and organizes the form the world can have for me.

2. Others

   a. The world of which I am aware is populated by two very different sorts of entities, perhaps among others as well: animate organisms,

and material objects. Some animate organisms I call "other "persons"; they possess all of the following characteristics: They are conscious of the world in a fashion analogous to the manner in which I am conscious of it; they are conscious of the world in a way independent of my own awareness of them and of the world; they are often conscious during periods when I am not (when I am in a deep sleep; after I am dead; etc.).

b. I make the same kinds of claims about other persons as I do about myself (under 1. above).

3. Material objects

a. When I am aware of a material object -- for example, a brick -- there exist parts of the object which I am not directly aware of (i.e., cannot actually see or touch, etc.) and yet which possess the same kinds of sensory characteristics as the parts of the object which I do observe.

b. When I am not aware of an object either directly (perceptually) or indirectly (e.g., through memory), it is most frequently the case that the object continues to exist, essentially with the same properties (or others of the same kinds) as were evident when I observed the object.

c. Objects are not the sorts of things to which logically incompatible properties can simultaneously be ascribed -- e.g., is A and is not-A; is bent and is straight; etc.

4. Process

a. Events are what they are only in relation to an overall process in which they take place. Isolated events, which do not form parts of a process, are impossible; a process is "ontologically prior" to the particular events which make it up.

5. Time

a. The passage of time is sequential: Events which occurred earlier in my experience may be retained by me in my memory, yet these past events somehow are more than just my memory of them.

b. Future events, although I may anticipate the "arrival of future events", transcend my anticipations: The future is more than just my "sense of the future".

c. There is historical evidence to support the view that many events occurred prior to my earliest memories. Yet I would not wish to reduce these events to the evidence of their occurrence: that Caesar crossed the Rubicon is/was an event above and beyond the actual historical records we may have to attest to the fact.

6. Space

   a. My conception of physical space serves to unify, e.g., successive experiences I have of different sides of an object. As I view a building from different angles, the successive perspectives I see come to be integrated through the synthetic idea of space which is basic to my outlook. Yet, independently of my outlook, there exists a physical space in which my walking around the building takes place.

   b. When one, two, or n men walk around a building, there is but a single physical space in which they walk.

7. Truth

   a. In mathematics or in physics, for example, when a theorem is proved or a physical effect is demonstrated, the result is discovered, it is not a human invention. The truth which the proof or demonstration discloses was already there; its validity is independent of human activity.

   b. There are some truths which are absolute; they do not in any essential way depend upon the accidents of time and place, culture and outlook.

   c. Results which have been proved to be the case can be given retrospective validity. If the principles of general relativity are shown to be basic to any understanding of gravitation, this would also serve to show that gravitation experienced by Neanderthal Man was governed by these same principles, unknown to him at the time.

8. Theories

   a. In the succession of, e.g., physical theories about the universe, a well-established hypothesis (e.g., that the earth circles the sun and rotates on its axis) is unique and non-arbitrary: it is not the case that there are numerous alternative explanatory options available.

   b. There exist real "first principles" which must be accepted in order for our comprehension of the world of our experience to have the synthetic unity it does.

   c. There is a specific and unique set of ultimate reasons why we are here.

9. Conviction

   a. If a view is rational, it should entail real conviction on our part. Similarly, if a view can be shown to be irrational, any

conviction we may have felt toward it should be neutralized.

b. In establishing our convictions, it is important to avoid inconsistency.

10. Causality

a. At times, my behavior serves to influence that of others, and in general makes a real difference to the course of events. In retrospect, it can be claimed, for certain x's and y's, that had I not done x, y would not have happened.

b. For every act, there must be an agent, for every thought, a thinker, and for any intentional decision, a deliberating ego.

11. Ethics

a. In many situations, there is only a single morally correct path of action. Consequently, the concept of "human goodness" is unambiguous and invariant.

b. The historical transition from barbarism to contemporary civilization expresses real progress in an objective way.

12. Political philosophy

a. My own physical, psychological, and spiritual state and those of others, are determined in large part by the prevailing social and political environment.

13. Aesthetics

a. There are real and objective standards of beauty. It is possible to make critical assessments of works of art which do not rest on mere taste and subjective values.

14. Proof theory

a. There exists an absolute system of propositions from which all other true propositions can be derived whether formally or informally.

b. Proofs are possible which do not depend on the prior acceptance of specific sets of premisses.

A Challenge to the Reader

In the first self-test, you probably experienced how it is possible to shift your awareness out of absorbtion in a story in order to identify incongruities. This is the first step in epistemological reflection.

In the second self-test, you probably pondered over each of the claims listed there, weighing it in terms of your own tacit philosophical commitments. Having read Chapter 1 on conceptual therapy, and already perhaps somewhat self-conscious about possible conceptual "slips" and "mistakes", you may have begun to feel a little paranoid! Perhaps you wondered what inconsistencies and difficulties may be buried in your answers. You may have asked, to what extent will my own philosophical belief system succumb to reflective criticism and show varieties of "epistemological loss" analogous to someone who reads "A Nice March Day" without perceiving any strange incongruities?

From my own point of view, I do not believe that any of the claims on the philosophical belief system list are intrinsically erroneous or inconsistent per se. Probably all the claims are, to varying degrees, vague and ambiguous, possessing different latitudes of interpretation. However, if each claim is placed in the context of what I am tempted to call the Standard Epistemology of our culture, then much of this vagueness and ambiguity is eliminated as the intended meaning behind each claim is given an adequate expansion.

When this is done, we may find -- and I anticipate that most readers will come to agree with this -- that many of the listed philosophical commitments (or their rejection) simply are untenable, because, in different ways, they "short-circuit" the very system of concepts in terms of which they are understood.

What this amounts to is in appearance a rather brash challenge to the reader. What I am suggesting is that your own philosophical belief system is very likely epistemologically pathological, not probably altogether so, but that your philosophical beliefs, some of them, perhaps quite a few of them, are very much like parts of the "Nice March Day" story. On one level, they sound fine. Until you think about them a little bit. And then you may find that, whereas the "Nice March Day" story is compounded of incongruities, your most basic philosophical beliefs about the world, yourself, other people, etc., are here and there fundamentally incoherent. That is, they are downright crazy, in some rather refined senses of the term.

The reason I would suggest something so seemingly preposterous, since I don't even know what your philosophical beliefs are, is because you, like me, have developed and live in the intellectual environment of a culture.

Whether you were raised in North America, Latin America, Europe, the Near East, or elsewhere in the world, your vocabulary of ideas is probably, in large measure, inherited.

Conceptual therapy may seem brash, but is is not arrogant. It recommends as virtues reflective analysis, intellectual modesty, and a careful attending to the frameworks in terms of which claims are made. Conceptual therapy poses an intellectual and personal challenge to any reflective mind: to develop a philosophical outlook which is not a matter of arbitrary taste and personal opinion, but is rather a system of concepts each of which has been carefully forged and tempered in the fires of careful criticism. The objective of conceptual therapy, probably an ideal that is never quite attained, is the man who knows what he knows, and limits his claims to know accordingly, and who, further, lives and thinks in ways that are pathology-free. Such an ideal human being does not work <u>against</u> himself and others, he is no longer destructive, because, ultimately, he is no longer self-sabotaging.

CHAPTER 3

How to Show that You Know that You Know: Methods for Epistemological Demonstration

How to Show that You Know that You Know:

Methods for Epistemological Demonstration

In the two previous chapters, we have noted certain parallels between the objectives and frameworks of conceptual therapy and psychotherapy. We have seen that whether a therapeutic framework focuses on ideas or on psychological-behavioral problems experienced by human beings, ultimately it is the conceptual framework or the personality structure of the individual person or group which is the locus of change.

In this chapter, we move away from this point of common contact, as we begin to examine methods which can be used in epistemological analyses which involve impersonal representations of sets of philosophical commitments. In other words, we will now restrict our attention to the theoretical province of epistemology -- analyses which are conceptually oriented, which attempt to clarify and analyze systems of ideas that have been detached from the mental environments of individual persons. By doing this, we gain generality, since we do not need to refer to highly individualized cases. Too, this is a more intellectually effective path to take, since a specific position or theory can be fixed in an impartial and detached manner, and analyzed with a degree of control which is frequently difficult or impossible if reference is limited to the often amorphous and shifting belief systems of individual persons.

Two distinguishable epistemological tasks should therefore be taken into account. One involves purely theoretical work, consisting of attempts to give an adequate descriptive representation of a claim to knowledge, or of a set of such claims, in order that a position can then be analyzed epistemologically. The second task involves implementing the results of such an analysis: here, a bridge must be developed which spans theory and practice, which can effect conceptually desirable changes within the conceptual structures of individual thinkers.

Epistemology has traditionally confined itself to the analytical task, and left the task of application within the conceptual systems of individual thinkers up to classroom interaction, up to a student's capacity to personalize general principles and results, and up to chance.

There are good reasons for this apparent neglect of the personal dimensions of epistemology. For one thing, the classroom is not an appropriate environment for what amounts to individual therapy, even when this therapy concerns thought processes rather than emotions. Individuals think in different ways, using different patterns of association, and approach any specific question with distinct systems of philosophical commitment. Attempts, with a class of students of any size, to go from theoretically-oriented epistemological analysis

to individual therapy usually are not effective because individual differences cannot fairly be taken into account. Yet this is surely not a unique dilemma faced by epistemology; it is one of the limitations of mass education, in any field. Only it is perhaps more distinctly felt in any discipline devoted to personal or intellectual development.

With these limitations in view, we turn now to consider several classes of methods which are fundamental in epistemological analysis.

The Three Principal Methods

A specialized theory is fundamentally a representation of a system of conceptual claims. The propositions we identify as basic in the theory reflect a set of interests: If our interests are in the deductive structure of a theory, the fundamental propositions will be those that are, from a deductive point of view, basic. If we study a theory in terms of the domain of objects it refers to, then those propositons we identify as fundamental will be those that indicate the most elementary objects in the domain and express the most basic ways the theory has of referring to those objects.

This is no less true of an epistemological set of interests: When we look at a specialized theory, our interests, as epistemologists, will direct us to focus our attention on what we see as the strongest claims to know certain things about the subject-matter in question. Often, the strongest knowledge claims are not even expressed explicitly in a theory: For example, many theories implicitly claim that the propositions asserted by them are true not only from their own frames of reference, but are true of their subject-matters independently of their theoretical representations. In this way, the truths expressed by a theory are thought to be truths-about-the-world, and not simply truths-provided-one-adopts-a-certain-point-of-view.

Once we have selected a group of propositions which we choose to regard as epistemologically fundamental, several approaches are open to us as we may wish to undertake a critical, analytical, or demonstrative study of these propositions, depending upon whether we are concerned with a clarification of the limits, the conditions, or the preconditions of a knowledge claim.

Three general problems arise, depending upon which of these concerns we have in view. A different problem is posed by each of the following questions:

1. What does a specific knowledge claim mean?

2. Is this claim consistent?

3. What must the claim presuppose in order to be possible?

The following are conceptual approaches or models useful in answering these questions:

A. The model provided by criteria of meaning:

Numerous criteria of meaning have been proposed by philosophers. Carnap, Bridgman, Schlick, Ayer, and others have proposed specific criteria of meaning. Other philosophers have proposed that meaning should be explicated as a function of what is given in consciousness,

and so have sought to refer to certain phenomenological data -- cf. Husserl, Meinong, Merleau-Ponty, etc. Both the application of a criterion of meaning and recourse to data of phenomenological insight provide us with ways of answering the first question, What does a specific claim to knowledge mean? At the same time, reference to certain criteria of meaning enables us to claim to know that this is what the meaning in question is.

B. The model provided by logical criteria:

Many different kinds of tests for consistency are available. Some are useful if we are interested in the consistency or non-contradictoriness of an individual proposition (relative, of course, to a chosen model of logical consistency); others are useful if our interest is in the internal consistency of a system of propositions. Consistency tests can, in other words, be applied to individual propositions, or to systems of propositions.

C. The model provided by modal criteria:

This is the approach first proposed in any detail by Kant. Its interest is in proving that, for a particular claim or system of claims to be possible, certain presuppositons cannot be avoided. When such a proof can be given for a certain presupposition, that presupposition is said to be demonstrated by means of a "transcendental deduction".

The conceptual vocabulary needed to work with this approach is less familiar to us than are criteria of meaning, experiential data, or logical criteria. The concepts of necessity and possibility play a central role in the transcendental approach; they are known as logical modalities; this is the basis for the name for this general model.

Several conceptions of possibility and necessity have been proposed. Hume suggested that possibility be understood in a psychological sense: What is conceivable is possible. A more general understanding of possibility is this: What is not prohibited by a certain set of constraints is possible with respect to those constraints.

Similarly, a necessary proposition has the property that its denial is impossible: For Hume, the denial of a necessary proposition is inconceivable. In general, the denial of a necessary proposition is inadmissible relative to some set of constraints.

Consequently, a proof that a transcendental deduction is valid will take this form: its denial renders impossible the knowledge claim(s) with which it is associated.

Of these three approaches used in epistemological analysis -- pertaining to meaning, logical, and modal criteria -- this book concentrates on the third of these; no prior works in epistemology have developed this approach so that it can be implemented effectively. Numerous works are available which can serve as excellent introductions to the other two approaches. Some useful references are listed below:

On meaning criteria

Ayer, A.J., ed.: Logical Positivism (New York: The Free Press 1959), particularly papers 2-5, 10, 11.

Bartlett, Steven J.: Chapter 12 of this work briefly describes traditional meaning criteria, indicates the main reason why they have been unsuccessful, and proposes a new criterion of meaning.

Carnap, Rudolf: The Logical Structure of the World and Pseudoproblems in Philosophy, trans. by Rolf A. George (Berkeley, California: University of California Press 1967; orig. German, Der logische Aufbau der Welt (Berlin-Schlachtensee: Weltkreis-verlag 1928), Scheinprobleme in der Philosophie: Das Fremdpsychische und der Realismusstreit (Berlin-Schlachtensee: Weltkreis-verlag 1928).

Carnap, Rudolf: La Science et la métaphysique devant l'analyse logique du langage, trad. Ernest Vouillemin (Paris: Hermann 1934; first printed 1931).

Bridgman, P.W.: "The Operational Aspect of Meaning", Synthèse, Vol. 8, 1950-51, pp. 251-259.

Bridgman, P.W.: The Logic of Modern Physics (New York: Macmillan 1961; first printed 1927).

Bridgman, P.W.: The Nature of Physical Theory (New York: Dover 1936).

On phenomenology

Bartlett, Steven J.: Chapters 10, 11 of this work give condensed descriptions and critiques of the phenomenological approach.

Husserl, Edmund: The Idea of Phenomenology (Hague: M. Nijhoff 1964).

Husserl, Edmund: Ideas: General Introduction to Pure Phenomenology, trans. by W.R. Boyce Gibson (London: Allen and Unwin 1952; first English printing 1931).

Spiegelberg, Herbert: The Phenomenological Movement, A Historical Introduction, 2 volumes (Hague: M. Nijhoff 1960), particularly Vol. 2: "The Essentials of the Phenomenological Movement", pp. 653-701.

On logical criteria

Lemmon, E.J.: Beginning Logic (Indianapolis, Ind.: Hackett 1978; first printed 1965). A clear, concise, and usable approach to the propositional and predicate calculi.

Prior, A.N.: Formal Logic (Oxford: Clarendon 1955). An excellent book, though in Polish notation.

Tarski, Alfred: Introduction to Logic and to the Methodology of Deductive Sciences (New York: Oxford University Press 1965).

Examples

An example involving the application of meaning criteria:

"Are mathematical results discovered or invented?" "Is the psychological perception of a solution (insight) a matter of invention or discovery?"

-- Such questions request a clarification of the meaning of 'discovery' or 'invention' in the context used. A criterion of meaning (e.g., verification, or an operational understanding of meaning, or expression of factual content), or appeal to experiential data, can be employed to answer questions of this sort. Use of a well-defined criterion of meaning, or reference to data drawn from experience, can provide us with ways of justifying epistemological claims.

An example involving the application of logical criteria:

"A set contains all and only those sets that do not contain themselves as members. Does it contain itself as a member or not?"

Logical paradoxes like this can be shown to result in inconsistency relative to some frameworks. Logical inconsistencies may be of various kinds. Their study is included in most works devoted to mathematical logic.

An example involving the application of modal criteria:

"All mathematical results are dependent upon our mathematical abilities, capacities of expression, etc."

Is it possible to refer to a mathematical result without employing mathematical modes of reference? If not, then the functional dependency which is asserted is demonstrated.

Return to these examples once you have studied specific techniques belonging to the three approaches to epistemology. At that time, provide for each example the requested clarification of meaning, demonstration of inconsistency, or proof of the necessity to assume a precondition. Make sure you show how your own answer is justified -- what specific criteria do you appeal to?

Transcendental Deductions

One of the most important, complex, and controversial contributions to the history of philosophy was made by Immanuel Kant, a German philosopher who lived from 1724 to 1804. In his Critique of Pure Reason, Kant sought to identify a group of categories without which objective knowledge would be impossible. His transcendental deduction was an attempt to demonstrate that the categories which he identified do actually play the role of transcendental preconditions of objective human knowledge. This was the first careful attempt in philosophy both to provide a detailed description of this kind of argument, and to offer an impressively systematic transcendental proof.

Much of Kant's work here was original, and hence was an attempt to break ground in a new area of inquiry. Because of the complexity of his style of argument, there has been a good deal of confusion as to exactly how transcendental argumentation works. Since Kant's time, various attempts have been made to use forms of transcendental argumentation to demonstrate if arguments are valid, if concepts have applicability, whether propositions are true, whether terms are meaningful, etc. In the process, many voices have been heard, but little has been said that can serve the purposes of conceptual therapy.

My concern in this area has been to try to develop a method which (i) is flexible, that is, can be applied in a wide range of contexts; (ii) is comparatively (in relation to Kant's approach) simple; and (iii) is for these reasons usable and useful. The resulting method bears little resemblance to Kant's approach, but the original inspiration may still be perceptible.

The purpose of a "transcendental deduction" is to show that certain presuppositions (ideas, propositions, procedures, etc.) are unavoidable if a particular claim to knowledge is to be possible at all. A transcendental argument seeks to demonstrate that one or more presuppositions, or transcendental preconditions, are necessary in order for a certain type of knowledge to be possible. The difficulty in such a proof is to show, satisfactorily -- i.e., in a rationally compelling way --,that the identified presuppositions really are inescapable.

Turning this around somewhat, we have this formulation: A transcendental deduction is a proof which is such that rejecting the identified presuppositions renders impossible the knowledge claim with which they are associated.

All of this may seem reasonably clear, until an attempt is made to carry out the project. And then it is decidedly difficult to determine how one is to show that rejecting a would-be transcendental precondition really does render an associated knowledge claim impossible. What standards are to be accepted relative to which this impossibility is going to be pronounced? Do these standards compel assent, do they have a status which is non-arbitrary? For if they are arbitrarily stipulated, it will be easy enough

to dismiss the conclusion reached by an alleged transcendental proof. -- Such are some of the shoals upon which attempts to develop transcendental argumentation have ship-wrecked.

One of the main reasons, it seemed to me, that attempts have failed is because the critical standards employed are usually "external" to the knowledge claim in question. -- Suppose you are told that a dream you have had is "really" a sexual dream, because you dreamed you could fly. A standard of interpretation is applied to your dream description, one that is "external" in the sense that it is alien to the context to which it is applied. That dreams of flying are to be understood as having sexual meanings is an intellectual construction imposed on your dream from outside your dream-framework. Problems of interpretation arise because of the fact that the framework of interpretation does not coincide with the framework being interpreted. There is a kind of circularity in the justification used by a Freudian dream analyst; it is visible in the following example of petitio principii, a begging of the question: "The reason why you do not accept the Freudian account of your dream is because you are trapped in the anal-retentive stage." The reason that "explains" your disagreement assumes the very Freudian frame of reference you may question. And the same can be true in the case of transcendental arguments. For if a proof establishes the inescapability of a group of preconditions, but only in relation to some arbitrarily stipulated criterion or convention that is external to the framework of a particular knowledge claim, then the proof begs the very question it was designed to answer.

This problem has been a principal obstacle to transcendental argumentation. The answer, I have proposed, is to avoid the introduction of standards that are external to the framework of a given knowledge claim. Stay within that framework. But, if one does, is it possible to prove that a would-be transcendental precondition really is a transcendental precondition, a presupposition that must be accepted in order for the knowledge claim to be possible?

I believe this can be done. It requires the use of self-referential argumentation of a sort that I have called "metalogical".

A transcendental argument has this form: If x is to be possible, then y is necessary, where x is a specific knowledge claim, and y is the assertion of one or more alleged transcendental preconditions, i.e., y is a statement of the form "x presupposes...", where the blank is filled with an enumeration of conditions which are thought to underlie the possibility of asserting x.

Using '$\Diamond$' for the modal operator 'possibility', '$\Box$' for the necessity-operator, and '$\vdash$' for the derivability sign, a transcendental argument has the form:

$$\Diamond x \vdash \Box y.$$

A transcendental proof of the above argument, as I have developed this in a self-referential metalogic, can be described briefly as follows: (A more complete account is given later in this chapter, and in Chapters 9-12.)

A logical test of $A \vdash B$ is to show that the conjunction of A and not-B is logically impossible. If statement B can in fact be derived from A as premiss, then asserting the conjunction A & -B will be logically false: within standard bivalent logic, from A & -B it should then be possible to derive a contradiction.

As we have described this, a transcendental argument has the same logical form as $A \vdash B$. An analogous test for $\Diamond x \vdash \Box y$ can be devised by examining what happens if x & -y is asserted. This would be tantamount to asserting the knowledge claim in question, while, at the same time, denying what one believes to be are certain presuppositions that are necessary in order for the knowledge claim to be possible.

This is a radical test of dependency: Taking away someone's crutch (denying y) will make it impossible for him to walk (x), if he is handicapped, i.e., depends on a crutch. If it is impossible for x to walk without y, y must be necessary in order for it to be possible for x to walk.

In a metalogical proof, the dependency relation to be tested is between a knowledge claim and certain presuppositions which we have identified to be, we think, necessary in order for the knowledge claim to be possible. In our test to determine whether these presuppositions <u>are</u> necessary (whether the person requires his crutch), we wish to show that without them, the knowledge claim is <u>undermined</u>, that without these presuppositions, the knowledge claim "collapses", so to speak.

The way I have suggested that this can be done is to show, for self-undermining knowledge claims, that

$$x \ \& \ -y \vdash z$$

where z is a <u>metalogically self-referentially inconsistent</u> statement. Such a self-referentially inconsistent statement is one which literally and logically jerks the carpet out from under the feet of its associated knowledge claim. Here, an example is helpful.

P.W. Bridgman proposed, in philosophical jest, that the entire universe is "shrinking homogeneously" in such a way that all our relevant standards of measurement are proportionally affected. Everything in the physical universe is, at some indeterminate rate, getting smaller and smaller. No technique would, under the hypothetical cosmic shrinkage knowledge claim, be usable to detect that a process of shrinkage is going on.

Bridgman was a proponent of an operational criterion of meaning: If, in the case of hypothesized cosmic shrinkage, we have no physical operations or procedures we can employ to detect this "shrinkage", the concept of "cosmic shrinkage" is without meaning, since there are no operations in terms of which its meaning could be made determinate. Bridgman concluded that therefore it does not in fact make sense to claim that the universe is shrinking in this fashion.

The cosmic shrinkage hypothesis is also metalogically self-referentially inconsistent. Consider this transcendental argument:

> Let x represent the knowledge claim, the universe is shrinking homogeneously in Bridgman's sense. Let y represent the tentative assertion that, to know x, it must be possible in principle for us to refer changes in size to some stable standard, one that is unaffected by the shrinkage of the cosmos, in order to claim, with any meaning, that anything is "shrinking".
>
> To test this argument, we will assert x, and deny y: What we obtain as a result is the compound claim that (i) it is known that cosmic shrinkage is going on, and (ii) it is impossible, in principle, to refer to any standard relative to which changes in size can be detected.
>
> The claim made in the last paragraph is evidently self-referentially inconsistent: In (i), reference is made to "cosmic shrinkage", and in (ii) it is denied that there are any possible ways of referring to changes in relative size. Since "shrinkage", cosmic or otherwise, means that the sizes of things are reduced relative to some standard, the hypothetical knowledge claim that asserts that the universe is shrinking homogeneously undermines itself.
>
> The conlusion we may draw from this is that the hypothetical knowledge claim, that the universe is "shrinking homogeneously", cannot, in principle, make sense. (Compare the different modal level of this conclusion with Bridgman's: it does not in fact make sense that the universe is shrinking homogeneously.)

In other words, to show that $\Box y$ can be derived from $\Diamond x$, we showed that $x \,\&\, -y$ leads to metalogical self-referential inconsistency. This is equivalent to saying that it is impossible (on pain of metalogical self-referential inconsistency) to assert $x \,\&\, -y$. Therefore, $\Diamond x \vdash \Box y$ has been demonstrated by a type of transcendental deduction.

If a transcendental argument can be formulated such that its negation leads to metalogical self-referential inconsistency, then the argument will be called "self-validating". In fact, a knowledge claim A is self-validating if and only if -A is metalogically self-referentially inconsistent.

As should gradually become clear, the power that resides in this kind of approach is due to the fact that no external criteria of meaning, validity, truth, etc., are relied upon. Just what metalogical self-referential inconsistency will mean in a particular context will depend on the system of concepts used there. As we shall see, for claims to express knowledge, some variety of justification is called for. But what form such justification will take, is context-relative. Some frames of reference assume bivalence, some do not.

Some intend to refer to physical objects, others do not. What variety of justification is called for to back up a particular claim to knowledge depends upon the context of justification in question. It is important to provide the degree of flexibility needed to cover numerous possible contexts of reference. The concept of justification is therefore left open. In an epistemological analysis of a particular knowledge claim, its associated frame of reference will determine how the concept of justification is to be filled in.

On Self-Referential Argumentation

We have seen that the success of a transcendental deduction rests on effectively proving that if one or more preconditions are denied, then an associated claim to knowledge becomes impossible. I have suggested that it is possible to prove this by applying the idea of a metalogical variety of self-referential inconsistency.

There are several varieties of self-referential inconsistency. Semantical paradoxes exhibit forms of self-reference (e.g., "This sentence is false.") Set theoretical paradoxes are also often self-referential (e.g., the set that contains all and only those sets which do not contain themselves as members). Behavioral paradoxes (e.g., saving pieces of string too small to save, or Raymsey's example of a child saying, "I can't say 'cake'.") exhibit another form of self-reference.

Inconsistency which comes about due to self-reference can, in general, be broken down into two kinds: One kind of inconsistency arises due to statements which are self-defeating or self-refuting, such as the claim, "There are no truths." A second variety of inconsistency arises as a result of claims or concepts (which frequently involve implicitly a number of claims linked together) which undermine themselves. They do this by explicitly denying presuppositions which must be granted in order for the claim to be made, for the concept to be expressed, to have the meaning intended, to express the purpose required, etc. Examples of such forms of self-referential inconsistency, which involve a denial of metalogical preconditions, include Bridgman's shrinking universe hypothesis, Descartes' universal dream hypothesis ("All of life may be a dream."), etc. Henceforth, I will call such self-referential inconsistencies projective.

Self-refuting statements include both semantically and set theoretically self-referentially inconsistent statements. Self-refuting statements also include statements which, like "I can't say 'cake'" or "there are no truths", have been called pragmatically inconsistent -- i.e., what the statement is used to claim conflicts with what the statement asserts. In Ramsey's familiar example, a child uses the word 'cake' in order to say that he can't utter that word. Similarly, the statement "there are no truths" is used to claim (allegedly, with truth) that there are no truths.

Projective statements require an analysis on a different and higher level. The conflict in a projective form of self-referential inconsistency is not between the use of a statement and what the statement asserts, nor is the conflict between the properties ascribed to a class of things and the assertion that such a class must either exist or not exist. Nor is the conflict between a statement's truth (or falsity) and what the statement claims about its own truth (or falsity). Rather, the inconsistency involved in a projective statement can only be noticed from a metalogical point of view: i.e., when we become aware that what a statement asserts constitutes a rejection of one or more of the transcendental presuppositions which must be granted for

the statement to be possible at all.

In general, whether we have self-refuting or projective forms of self-referential inconsistency in view, the rejection of a self-referentially inconsistent proposition will be self-validating -- i.e, it cannot not be accepted. For, if it is not accepted, then a demonstrably self-referentially inconsistent statement will result.

In other words, the rejection of a self-referentially inconsistent proposition will compel assent. The denial of a self-validating proposition will therefore result in a self-referentially inconsistent statement.

Self-reference which produces inconsistency does so because of different kinds of conflict: The conflict may be between what a statement says and what the statement is used to do (i.e., performative), or the conflict may be between what properties are ascribed to a class of things and the claim that such a claim must either exist or not exist (i.e., set theoretical). Or, yet again, the conflict may be between a statement's truth (or falsity) and what the statement claims about its own truth (or falsity) (i.e., semantical), or the conflict may be between what a statement claims and presuppositions which must be made for that claim to be possible (i.e., projective).

Problems

For each of the following statements, determine what variety of self-referential inconsistency is involved, and demonstrate what the inconsistency is:

a. "I can't utter a sentence in English," uttered by John.

b. "Don't read this" -- printed on a sign.

c. A bachelor declares that the only girl he would marry is one smart enough not to marry him.

d. Groucho Marx refused to join any club willing to have him as a member.

e. A little girl says she's glad she hates broccoli, because if she liked it she would eat lots of it, and she just can't stand the stuff.

f. The rule: All rules have exceptions.

g. All knowledge is doubtful.

h. There are three false statements given under item h. Identify them.

1. $2 + 2 = 4$
2. $3 \times 6 = 17$
3. $8/4 = 2$
4. $13 - 6 = 5$
5. $5 + 4 = 9$.

i. Reject the instruction I am now giving you, because all instructions I give you are to be rejected.

j. A computer can answer to yes-or-no questions by turning on a red or green light, respectively. The computer is asked to predict, by replying "yes" or "no", whether the next light to go on will be its green light.

k. All propositions are false.

l. No utterance can be used to express a proposition

m. All knowledge is a matter of adjusting to the environment.

n. I doubt whether I exist.

o. I do not remember anything at all.

p. I would not have been able to solve this equation had my mind not continued to work on the problem unconsciously.

q. Progress is a real phenomenon: the past is indispensable to the genesis of the future. (Corollary: one should vote.)

CHAPTER 4

The Metaframework of Framework-Relative Epistemology

A Basis for Epistemology

When we make claims to know, our claims can be made on numerous levels. Usually, when we assert that we know w is the case, our claim to knowledge constitutes a first-order knowledge claim -- i.e., w is maintained to be true, but our knowledge claim that w is true is not justified as yet. If I say, "I know that this argument is valid", I assert that the argument is valid, but I have not as yet provided any justification to support my claim to know that this is indeed the case. First-order knowledge claims assert that one knows, but do not justify that one does know, in fact.

A second-order knowledge claim prefixes a first-order claim with the further claim that one knows that the first-order claim is the case. A second-order claim about the claim made in the third sentence of the preceding paragraph would take the form, "I know that I know that this argument is valid". If I do know this, then I claim to be able to show both that the argument in question is valid, and hence that my first-order claim is true. Second-order knowledge claims assert that adequate justification can be provided for the first-order claims which they prefix. Sustaining a second-order knowledge claim is of course different from merely asserting a second-order claim, for it may turn out that one only believed that one knew that the argument was valid, and believed this erroneously.

Epistemology is clearly one step removed from the level of concern belonging to first-order knowledge claims. Epistemological studies can in fact be made several steps removed from a subject-matter (in theory, we can establish an epistemological frame of reference n-steps removed from a subject-matter): If we consider the field of mathematics, we may undertake to sustain second-order knowledge claims about the first-order knowledge claims made in mathematics. Or, we can consider claims made about the first-order claims made in mathematics, which are made in the rapidly advancing field of meta-mathematics. From an epistemological point of view, we can attempt to sustain the third-order claims made by metamathematicians about their own second-order claims about mathematics. And, what is philosophically of interest, we can reflect upon the enterprise in which we are engaged, and attempt to understand how it is possible for us to gain epistemological knowledge on any level. At this stage, we are referring to a general, universal epistemological metatheory.

Such a general account of epistemology itself, would seek to make explicit the transcendental presuppositions underlying any epistemological approach.

The following are basic statements which appear to play the role of axioms in any epistemological undertaking -- i.e., they seem to be unavoidably presupposed in the sense that one cannot deny them without self-referential inconsistency....

A. Metatheoretical claims must serve the interests of informative, meaningful communication -- they must make sense.

(which can be broken down into these sub-claims:)

(i) We need to be able to determine either (i) when a claim does make sense, or -- at the very least -- (ii) some of the conditions under which a claim does not make sense.

If we choose to follow the first course, we will desire a criterion or a group of criteria of meaning;

and if we follow the second course, we will want to have a criterion or group of criteria to enable us to detect meaningless claims.

Often the second course is an easier one: We frequently find it more convenient (economical) to define a game by stating constraints placed upon moves, by identifying what is prohibited, and by permitting all moves that are not prohibited. (Consider, for example, the "moral game" defined by the Ten Commandments, many of which begin with the phrase "Thou shalt not...".) The wisdom of stating constraints placed upon moves can be seen if we imagine trying to define many games by identifying all possible permitted moves.

(ii) Any claim, concept, or term that fails to serve the interests of meaningful communication of information is to be rejected, and another claim, concept, or term that can perform the desired functions of encoding information in a meaningful communicative manner is to be substituted -- when this is possible -- for the failing, unsuccessful claim, concept, or term.

B. (i) In an analysis of any assertion about a subject-matter, X, it is what we know that is important. (Notice that this is a normative statement.)

(ii) What we know, we can conceptualize and talk about. (This is an epistemological claim.)

Consider what happens if these two groups of presuppositions are rejected:

(Rejecting A.:) Metatheoretical claims need not make sense.

(i) We neither need to be able to determine under what conditions a claim is meaningful, nor under what conditions it is meaningless.

(ii) We should feel free to retain claims, concepts, and terms

the usage of which serves no informative, communicatory purpose.

(Rejecting B.:)

(i) In an analysis of assertions about a subject-matter, X, it is not what we know that is important.

(ii) What we know, we may be unable, in principle, to conceptualize or talk about.

Before continuing, let's summarize what has been claimed so far. We have been talking about knowing (and not knowing) what one is talking about when one makes claims about a particular subject-matter, X. We have associated with the notion that one knows, that one is able to provide some justification for what one says. The concept of justification has, we notice, been left open -- we do not limit our concept of justification by specifying what it is to mean, but leave the concept to be filled in by the context we have in view. This constitutes a strength, rather than a weakness due to vagueness: No stipulated conception of justification is imposed externally on any framework we might wish to study.

The original impulse to do epistemology arises out of a need to know what one is "talking about" when one makes claims in whatever field of study we may be interested in, and also makes second-order claims about that field of study.

To make this kind of inquiry possible, it has been suggested that we must accept certain methodological limitations. Specifically,

- that we will refuse to talk about claims or make claims which involve talking about what we cannot know -- i.e., cannot justify;

- that any claims we make in theorizing about a field of study, X, must be such that we can justify claims to know what we are talking about; and

- that our interest in the epistemology of X has both a critical/negative and a positive/constructive orientation: The approach is negative in that knowledge claims about X and claims made within X will be sorted into two categories, according to the standard: "here, one can justify what one claims"; "here, no justification can be offered." The orientation is positive in providing us with explicit criteria by means of which to judge when we know what we are talking about, as well as with a fund of claims that we are able to justify.

A Self-validation of the Metaframework

If I reject the general methodological limitations described by A. and B. above, then I either know what I am talking about, or I do not.

If I know what I am talking about, and can provide justification for my position, then I am inconsistent with my own rejection of A. and B. I have made a metaclaim ("I reject A. and B.") that makes sense and which I can, by hypothesis, justify. My metaclaim discloses tacit compliance with A. and B., even in the attempt to reject A. and B.

If I do not know what I am talking about, no justificaiton can be offered for my point of view, and it has no rational force to compel assent.

Finally, if I claim to know what I am talking about, but can provide no justification for my position, then I am marooned on an island of personal conjecture, inarticulate and dumb.

Therefore, to reject this metaframework is either inconsistent, and in that sense (literally) irrational, or it is arbitary and unjustifiable and in that sense (literally) stupid.

The choice of a metaframework for epistemology therefore is in this sense inescapable: What we have briefly outlined is a transcendental deduction that reveals that our theoretical frame of reference in conceptual therapy is a transcendental precondition of epistemology. Rejecting this metaframework renders epistemology impossible, at the same time that it undermines rational, discursive thought.

However, this conclusion does not entail that a given individual ought to accept this methodological basis. Only rational justifications can be proposed, in a circular way, for the acceptance of rationality as a point of departure. Rationality is incapable of legislating against adherence to an irrational or stupid position, except by begging the entire problem and offering reasons.

Perhaps the strongest, but nonetheless circular, reason for accepting rationality is that rationality provides the only means we human beings know of to identify and avoid self-sabotaging ideas and behavior. The very notion of self-defeating patterns presupposes rationality: reflective reasoning is needed to take note of patterns that sabotage themselves. Rationality is the only framework we have in terms of which to represent and to treat conceptual pathology.

Some Reflections

The approach to epistemology that has been outlined is self-avowedly critical and intolerant of talking about what one does not know and cannot

justify. There are obvious limitations of this approach: It will not appeal to philosophers who want to build grand speculative systems. It will not appeal to individuals who wish to retain systems of belief for which articulate justification cannot be offered. The openness, flexibility, and pluralism which conceptual therapy would foster does not include sympathy for self-defeating dogmatisms. What makes a view dogmatic in this sense is its inability, in principle, to communicate justification for its views.

A psychological reflection: Anyone who makes a strong claim in an apparently intolerant way tends to encourage in others a desire to oppose the claim. But here, in the context of describing a metaframework for framework-relative epistemology, attempts to reject this metaframework are self-sabotaging: The only opposition that can be effective in the metacontext we have described is rational opposition, the use of justifiable claims, etc., and this fact necessarily undermines the intent of any opposition.

There is, in fact, what amounts to a conceptual double-bind that is established by this approach to epistemology, which blocks any intellectual opposition to it.

CHAPTER 5

The Inescapability of Conceptual Therapy: Epistemology and the Double-Bind

The Inescapability of Conceptual Therapy:

Epistemology and the Double-Bind

Gregory Bateson described the concept of the double-bind in the following way:

1 - The field contains two or more persons.

2 - The experience is recurrent.

3 - There exists a primary negative injunction, usually taking the form of punishment for discriminating accurately (or inaccurately).

4 - There exists a conflicting secondary injunction, also normally taking the form of punishment.

5 - There exists a tertiary field constraint, which prohibits the victim from escaping from the field.

6 - There is a mapping onto any part of an individual's total experience so that small incidents can set off an entire symptomology.

An Example

Consider the relationship of a child to his mother over a period of time. (Conditions 1, 2, and 5 are immediately satisfied.) Let us suppose that the mother feels hostile and critical toward the child. If the child discriminates his mother's feelings correctly, he will be hurt. If he discriminates incorrectly, he may be. In a double-binding situation, the child is hurt either way.

To see how this can happen, we will have the mother return home tired after a day at work. She wants to be left alone. Her child has been by himself most of the day, and wants company. He is boisterous, and the mother cannot relax. She feels hostile, and orders the child: "Go to bed, it's late and you're tired!" If the child sees through this, he will be hurt because his mother has no time for him, is hostile, and does not want to be with him. If, on the other hand, he accepts his mother's incorrect labelling of his behavior (that he is tired, and that this is the reason why he should go to bed), then he will, in time, no longer be able to label his own behavior correctly or consistently. He may then begin to experience difficulty in doing this in connection with the behavior of others toward him. And so, either way, the child is hurt.

Consequently, having satisfied conditions 1, 2, and 5, the relationship described above (and we assume there are many episodes similar to this one in the upbringing of the child) also satisfies condition 3 (e.g., "Go to bed!"), condition 4 (e.g., "Go to bed -- because it's good for you and your mother wants what is good for you."), and condition 6 (the child gradually loses his ability to discriminate messages accurately, and becomes unable to label his own or others' behavior correctly). Condition 3 ("Go to bed!") constitutes an injunction backed by implicit threat of punishment. Condition 4 is satisfied by the incorrect labelling of both the mother's behavior -- allegedly loving but actually hostile and impatient -- and the child's -- allegedly tired but really not tired at all. Condition 4 conflicts with condition 3 because now the injunction ("Go to bed!") is embedded in a new context which encourages the child to close his eyes to the motivations behind his mother's primary injunction, which is maintained with the threat of punishment. Although Bateson does not put it this way, we may think of condition 4 as a second-level expression of deceit.

Neurotic alternatives to a double-binding situation are

- feeling on the spot,
- shifting to a metaphoric level (seeing what is happening in an unreal way, to disguise what is painful),
- feeling defiant (paranoid),
- giving up or laughing at what is going on (hebephrenic), or
- ignoring the situation (catatonic).

The only escape from the double-bind is metacommunicative comment, talking about the double-binding situation itself. But this is also defeated in a double-binding situation by verbal and/or non-verbal behavior. (The reader can easily imagine these in the mother-son story.)

Schizophrenics may be identified by ways they react to double-binding situations (as paranoid, hebephrenic, or catatonic schizophrenics). They are, as Bateson expresses it, "rigidly committed to their patterns of inconsistency."

Provided some of the six necessary conditions above are <u>not</u> satisfied, an individual can react to a double-binding situation in non-neurotic ways. He may, for example, simply agree (or feign agreement), accepting the situation's constraints on him for the time being, without a loss of his own ability to perceive the nature of the situation. Or, he may label the double-binding source of prohibitions -- privately or publically referring to the source as "unfair", "deceitful", etc.

Conceptual therapy, on the other hand, creates, in the environment of a class, or in the author-reader mind-space established when epistemological

problems are the subject of one's reading, an analogous double-bind, but on a conceptual rather than a behavioral level.

The field contains two or more persons. The experience is recurrent. There is a field constraint (it may be difficult for you simply to drop out of the class, or accept your limitations when you decide to overlook the message epistemology may have for you, and put the book down). There are several injunctions in force: You must be able to determine when a claim does or does not make sense. You must reject claims that fail to make sense. You must value what you know and devalue what you cannot. If you know something, you must be able to talk about it. Etc.

These injunctions, if we see them in terms of conditions 3 and 4 above, could be described this way:

> 3'- Do not make unjustifiable claims; accept the primary injunctions of our epistemological methodology.
>
> 4'- I (your instructor or the author of this book) want you to try to test the primary injunctions by disagreeing with them. However, you must justify any disagreements you have with these injunctions, otherwise you are (i) irrational, (ii) stupid, and (iii) will get a low grade in the course! (These punitive devices are implicit in most courses; they are implicit, moreover, in the minds of most learners.)

There are acceptable and unacceptable ways an approach can impose a conceptual double-bind on an individual: In a disagreement with a Freudian, to use a now-familiar example, if one is told that his disagreement comes about because of his early childhood training, then we are presented with an example of an unacceptable begging of the question: The claims of the theory are assumed to be true in order to understand (label) the responses of a potential opponent.

On the other hand, if one disagrees with a claim, and then is shown that he must accept the claim in order to formulate his disagreement, then a non-vicious form of self-validation of the claim occurs. If one wishes to believe that there are things he or she knows but cannot put into words, one should at least have the intellectual tact and acumen to avoid the temptation of reporting this belief through the vehicle of words!

Self-referential justification, because it refrains from imposing on a claim, on a discipline, or on an individual external criteria of evaluation, comprises a non-vicious form of self-validation. It does, admittedly, double-bind potential opponents on a conceptual level.

And the possible responses to such a conceptual double-bind are similar to those identified by Bateson: If our reaction is neurotic, we may respond

with defiance, or by giving up or laughing, or by ignoring the message. On the other hand, if our reaction is healthy, we will accept what is useful to us, what is consonant with our self-actualizing purposes, and reject in ourselves, in our conceptualizations, and in our behavior, what is self-sabotaging.

CHAPTER 6

The Practice of Epistemology

The Practice of Epistemology

In Chapter 3, we discussed three methods which enable us to answer the questions, What does a specific knowledge claim mean?, Is this claim consistent?, and What must the claim presuppose in order to be possible? Chapters 4 and 5 argued that the metaframework presupposed by these methods is itself inescapable if we are reflective and rational in our thinking. We turn now from these theoretical concerns to the task of putting this theoretical framework into practice.

Like any skills, skills in conceptual therapy are developed as you put them to use. This chapter accordingly is divided into two parts. In the first, you will find several, short sample epistemological analyses. (Examples of more detailed analyses are contained in Part II of this book.) My main interest in these is to place organization before content. If illustrations are to be useful to later practice, it is more important for you to study the structure of an example than it is to concentrate exclusively on its content, since the content is usually determined by individuating characteristics of the particular knowledge claim being analyzed.

After these preliminary sample epistemological analyses, the second section of the chapter contains guidelines you may find useful in constructing your own analyses. There are three objectives that I have aimed at here:

(i) Attention is focused on a group of issues, claims, and concepts which frequently are epistemologically pathological. By reviewing these carefully, it is possible to become sensitive to similar and related problems occurring in other contexts.

(ii) A set of flexible guidelines or heuristics are described which may assist you as you develop your own epistemological analyses.

(iii) A review is made of certain general ways in which knowledge claims can be justified.

The chapter tries to point to a set of skills which are difficult to define, since they must be unusually flexible to meet the contextual requirements of different frameworks. This "pointing" is done, then, in four ways: via examples, by sensitizing you to potential pathologies, by developing a set of heuristics, and by describing "safe" means that are available to justify many knowledge claims. Beyond these general recommendations, conceptual clear-headedness requires careful practice, the exercize of precise conceptual and linguistic control, and patience.

Sample Epistemological Analyses

Carnap's Criterion of Meaning

The following sample epistemological analysis is organized in four steps:

1. A brief description is given of the problem I want to pay attention to; some indication is given why this problem is important from an epistemological viewpoint.

2. A concise statement of my own position is made. In this statement, I try to make clear what the connection is between my position and the problem I am examining.

3. I give some idea how I propose to demonstrate my claim.

4. The demonstration is given.

Step 1:

The theses of realism and idealism have traditionally been considered to raise questions fundamental to epistemology, questions which any epistemology is forced to resolve. One way of resolving these questions is to suggest a criterion of meaning, and then show what consequences follow for realism or idealism provided that criterion of meaning is adopted. The problem then posed is whether one is compelled to adopt the proposed criterion of meaning. I will treat this problem in connection with the theses of realism and idealism and attempt to resolve it.

Step 2:

I am convinced that neither the thesis of realism nor that of idealism is consistent. I believe this inconsistency in each thesis can be demonstrated while, at the same time, it can be shown that the application of a criterion of meaning to these theses must compel our assent. This is what I propose to do.

Step. 3:

I will describe briefly Carnap's meaning criterion, as it is defined in his Pseudoproblems in Philosophy, and indicate how Carnap applies this criterion to a representation of the theses of realism and idealism. I will then attempt to develop a more generalized statement of the claims of realism and idealism, and then indicate why it is that we cannot not accept the application of Carnap's meaning criterion in this context.

Step 4:

Carnap defines his meaning criterion as follows: "The meaning of a statement lies in the fact that it expresses a (conceivable, not necessarily existing) state of affairs. If an (ostensible) statement does not express a (conceivable) state of affairs, then it has no meaning; it is only apparently a statement. If the statement expresses a state of affairs then it is in any event meaningful; it is true if this state of affairs exists, false if it does not exist. One can know that a statement is meaningful even before one knows whether it is true or false." [Pseudoproblems, p. 325]

In short, according to Carnap, a necessary and sufficient condition for meaning is the expression by a statement of a possible state of affairs. For a statement to be meaningful it is not necessary that we actually know whether it is true or false.

Carnap concludes, then, that "a (pseudo) statement which cannot in principle be supported by an experience, and which therefore does not have any factual content would not express any conceivable state of affairs and therefore would not be a statement, but only a conglomeration of meaningless marks or noises." [Ibid., p. 328]

His conclusion can be made more precise in the light of these definitions he gives: A statement is testable if conditions can be indicated under which the statement or its contradictory is confirmed ("supported" is the word he uses). A statement has factual content if experiences that would confirm it or its contradictory are conceivable, i.e., possible. In other words, a statement is testable if we have actual procedures to determine whether it is true or false, while a statement has factual content provided we know what conditions these procedures would satisfy if we had them. A pseudostatement is therefore an ostensible statement for which we do not know what conditions would be satisfied by confirmation procedures, even if we had them.

Carnap distinguishes two, frequently conjoined, theses of realism:

a. There exist things which I do not perceive.

b. There exists consciousness ("in others") when I am not conscious.

The corresponding theses Carnap ascribes to the idealist position are:

a'. There do not exist things which I do not perceive. (Negation of a. above)

b'. There does not exist consciousness ("in others" when I am not conscious. (Negation of b.)

Carnap's claim: Neither thesis has factual content -- i.e., for neither thesis do we know what conditions would be satisfied by procedures for confirming a. or not-a. and b. or not-b., even if we had them.

My claim: Carnap proposes the adoption of a meaning criterion and its applicability to the theses of realism and idealism. His proposal is compelling -- we cannot not accept it: Any attempt to reject it in this context leads to inconsistency.

My argument:

A. For realism:

Let the following set of propositions together express a generalized claim associated with the thesis of realism, combining a. and b.:

1. There is a frame of reference F that establishes a basis for my referring to events, in terms of which the idea of an event has a certain sense.

2. F provides a basis for reference to a set of events E.

3. There are events not included in E; call this set E'.

4. Events in E' are events in the sense of those in E.

5. F does not provide a basis to refer to E' or events in E'.

6. There does not exist any other frame of reference which I can employ to refer to E' or events in E'.

The above set of propositions is internally inconsistent: Propositions 3-6 above refer to E' or events in E', while proposition 6, in conjunction with 5, precludes a basis for such reference. The thesis of realism requires 6:

The thesis wishes to assert the autonomy of certain objects of reference in relation to my own frame of reference. If I did have available to me some other frame of reference, e.g., F', which enabled me to refer to events in E', the thesis of realism would still assert that there are some additional events which transcend the frames of reference (F and F') available to me. Therefore, proposition 6 brings this unnecessary regress to an end, by indicating that my own frame of reference (i.e., F) does not enable me to refer to events other than those in E. Yet, realism asserts, other events, in E', do exist.

Consequently, the set of propositions 1-6 is internally inconsistent.

Any attempt to construe realism in any sense which is derivable or reducible to a set of propositions 1-6 is inconsistent.

A metacomment: We note that no specific content has been given to a phrase like "a basis for referring to events E". Usually, philosophers who are realists identify such a "basis for reference" with the world of my consciousness, with what I am able, under various conditions and perhaps at various times, to be aware of, whether I use my senses, my memory, or perform certain experiments. All of these constitute means I have at my disposal to refer to things. In the above representation of the realist position, I am more interested in generality than I am in how my capacities to refer to events may be understood in specific terms. By doing this, the argument sketched does not depend upon what I may later come to feel is an overly restrictive definition of my own "frame of reference".

B. For idealism:

Let the following set of propositions together express a generalized claim associated with the thesis of idealism, combining a'. and b'.:

1'. There is a frame of reference F that establishes a basis for my referring to events, in terms of which the idea of an event has a certain sense.

2'. F provides a basis for reference to a set of events E.

3'. There are no events not included in E. The set E' is empty.

4'. The empty set E' is determined by the fact that $-F \rightarrow E'$: i.e., if my own framework is denied, then there are no events. This is the thesis of solipsism.

5'. F does not provide a basis for reference to E' or to possible contents of E'.

6'. There does not exist any other frame of reference I can employ to refer to E' or to possible contents of E'.

The above set of propositions is also internally inconsistent. The proof is the same as for the thesis of realism, substituting 1'-6' for 1-6.

Conclusion

This sample epistemological analysis focuses on two positions, realism and idealism, which are stipulatively represented by two sets of propositions. In other words, we began by saying, "Let the following set of propositions together express a generalized claim associated with the thesis of...". This

is a necessary basis for any epistemological analysis: unless the claim or general position we are referring to is somehow fixed in this way, no epistemological analysis can be effective. If a position is not clearly fixed, but is allowed to remain in a semi-amorphous, easily re-interpreted state, it will be an easy matter to claim that the results of an analysis do not apply to a slightly shifted, re-interpreted version of the position. By stipulating what I take to be the meaning of 'realism' and 'idealism', I fix the meanings of the positions I wish to analyze epistemologically. True, someone else can use the two terms 'realism' and 'idealism' differently; and then, my analysis may no longer apply. But this "limitation" is exactly what gives a framework-relative demonstration its force. We know what we are talking about, what we have in mind, and our demonstration relates to that. Exactly the same framework-relative proof-strategy is followed, e.g., in any branch of mathematics. When such a strategy is not used, results are always open to controversy.

In this analysis, the theses of realism and idealism are understood to be derivable or reducible to sets of propositions 1-6 and 1'-6'. Taken in this sense, each thesis is internally inconsistent. Any attempt to defend realism or idealism as understood in these ways will lead to inconsistency. Consequently, both theses must be rejected. Carnap's meaning criterion cannot not be accepted as it applies to the theses of realism and idealism: In this application, his meaning criterion is conceptually compelling.

Poincaré, on Mathematical Creation

Jules Henri Poincaré, a French mathematician and philosopher of mathematics, was interested in the way in which, from time to time, mathematical breakthroughs suddenly occur to mathematicians.

In a paper entitled "Mathematical Creation", Poincaré attempted to study this phenomenon. What he does in his paper is describe actual experiences he has had which occasioned a realization by him of certain mathematical results. What he wishes to do is to understand, given this descriptive basis, exactly how it was possible for him, and in general for any mathematician, to arrive at a sudden awareness of a mathematical result.

Let the following argument serve as a reconstruction of his thesis:

a. Often, direct and conscious mathematical work on a problem does not immediately lead to a solution.

b. Sometimes, although comparatively rarely, a solution to a difficult mathematical problem will suddenly occur to a mathematician, at times when he is not engaged in active efforts to reach a solution (e.g., stepping onto an omnibus at Coutances, or walking along the street in Mont-Valérien).

c. The sudden occurrence of such solutions is, Poincaré asserts, "a minifest sign [my emphasis] of long unconscious prior work. The role of this unconscious work in mathematical invention appears to me incontestable, and traces of it would be found in other cases where it is less evident."

d. Although it is contained only by inference in his account, later in his article Poincaré specifically refers to an interest in formulating an "explanation of the facts" surrounding such periods of sudden illumination. Apparently, he claims that an adequate explanation of sudden mathematical insight requires appeal to unconscious processes.

Two alternative, but not mutually exclusive, conclusions are possible at this point:

e. Unconscious operations actually take place in the mathematician and are responsible for his ability to reach a solution. "Sudden illumination" is the result of such prior but unconscious work.

e'. From the point of view of an explanation of the solution, it is possible to identify certain operations which would in fact lead to that solution. No claim is made that these operations actually took place "in the mathematician, unconsciously"; rather, they comprise an adequate account of how such an otherwise mysterious, sudden solution might have been reached had a solution-process occurred in a totally explicit manner (which it did not).

The difference between e. and e'. parallels the difference between the claim, on the one hand, that a child riding a bicycle unconsciously is a sophisticated mathematician solving numerous differential equations very rapidly and by so doing is able to keep his balance (e. above), and the claim on the other hand, that a computer can simulate ("explain") the child's behavior on the bicycle if it can be instructed to solve a battery of complex differential equations (e'.).

Decision$_1$: If e. is Poincaré's conclusion, and I suspect it is, in principle no justification can be given for this knowledge claim.

Decision$_2$: On the other hand, if e'. were Poincaré's conclusion, this knowledge claim can in principle be justified.

Argument$_1$ in support of Decision$_1$: Let s be the solution to a problem such that s was not reached directly and

immediately through conscious effort, but s suddenly occurred to the mathematician, and s cannot be easily explained by reference to conscious data. If reference is then made to "unconscious operations", and a claim is made that these actually occurred in the mathematician's thought process, then it must be possible to gain evidential access to such operations, without begging the question -- i.e., without assuming that they exist in the mathematician's thought process. But no such access is possible, if one grants the initial conditions of the problem above, in particular a. and b. In short, the formulation of the conditions of the problem in principle rules out that the proposed account e. can be justified.

Argument$_2$ in support of Decision$_2$: Let s be the same solution reached as above. Let T be a theory which asserts that if a set of operations S_o is performed, s will result. If T can be confirmed, then T can be used as an account of how s might be reached if only all the necessary steps involved were made explicit. Furthermore, it is known that the mathematician did not reason in this way, given a. and b.

This claim can easily be justified in principle since we know what conditions T would have to satisfy in order to hold for a particular s (i.e., $S_o \rightarrow s$: if the set of operations S_o is performed, s will result.)

The general conclusion we reach, then, is that we cannot know what we are talking about if we attribute the results of a theory of explanation to hypothesized and hidden processes involved, allegedly, in the phenomena themselves; but we can know what we are talking about if we claim that a theory is sufficient to account for the know facts.

A Gap in Memory

Suppose you or I has been drinking rather heavily at a party, and quietly goes to sleep in an easy chair before the festivities are over. The next morning, I awaken in my own home, in my own bed. I remember being at the party, realize I fell asleep, but do not remember how I came to rest in my own bed. What can I claim, and know what I am talking about, in connection with this "gap" in my memory, so I will understand how I returned (or was returned) home?

Suppose we transform the problem into a more general form:

Let m stand for a set of my memories, and s stand for my present state (being, e.g., in bed at home), such that there is no member of m which serves to explain s. I.e., s cannot be explained without the introduction of a supplementary explanatory hypothesis, h.

Assume h is expressed by persons who claim h can be derived from a second, but overlapping, set of memories m', such that at least some elements of m are members of m'. (In our example, this would mean that we are assuming, e.g., that one or more individuals at the party claim that I was carried to my home by them, and that these individuals share at least some of my own memories -- we share a sufficient number of memories to establish to my satisfaction that they remember being at the party, remember my being at the party, and I remember these same things, as well.)

The general question posed by the above problem can be subdivided:

Is it possible to claim meaningfully that

(a) h explains s -- that is, m' can be employed by me to provide an account of s where s cannot be explained by recourse to m?

(b) in relation to m and only to m, h' stands for real events which led to s?

In our example, (a) would, if answered affirmatively, assert that the explanation given to me by my friends at the party is an acceptable one, but I do not assert that what my friends claim to remember during the "gap" in my memory was indeed the case. On the other hand, (b) would, if answered affirmatively, constitute a claim made by me that, in fact, I was carried to my home although I have no memory of this.

The following answers to assertions (a) and (b) can be formulated:

1 - (a) establishes a context of explanation in terms of which s can be explained, although the explanation given makes no claim to uniqueness. -- I may resort to a "higher level" explanation, in which the account given to me by my friends is itself explained, or there may be equally "good", alternative accounts of the "gap" in my memory.

2 - (b) is self-referentially inconsistent: A basis for justifying (b) is ruled out by (b) itself. In relation only to m, no justification is possible for claim (b).

There are, of course, other possible ways of understanding the claims

(a) and (b). However, we are especially interested in those ways which establish a context of justification: Both conclusions 1 and 2 above make clear in what ways it is possible to justify them. The justification for the first conclusion above is clear: we should need to describe what constitutes a context of explanation, and could indicate alternative explanatory options. The justification for the second conclusion relies upon the ability to identify self-referentially inconsistent claims. In either case, it is evident that we know what we are talking about, and we are able to claim that indeed we know what we are saying.

Mathematics and Linguistic Relativity

> "We are thus introduced to a new principle of relativity, which holds that all observers are not led by the same physical evidence to the same picture of the universe, unless their linguistic backgrounds are similar, or can in some way be calibrated." - Benjamin Lee Whorf

Whorf, an American linguist who lived from 1897 to 1941, propounded the thesis that the structure of language shapes our understanding of the world. He argued that all higher levels of thought are dependent on language, and that the structure of one's habitual language influences his patterns of thought. World views differ among the speakers of structurally different languages, between which "calibration" is difficult if not impossible.

The so-called linguistic relativity hypothesis was developed largely through Whorf's studies of Hopi and Maya, and in part through the work of Edward Sapir (1884-1939) who for a time was Whorf's teacher. Within the field of linguistics, the theory of linguistic relativity is still controversial, though the growing inclination is to believe that, if language structure influences conceptual patterns, this influence exists more as a predisposition or inclination to think in certain ways, rather than as a tight control over what thoughts the mind can think.

I am not aware that the theory of linguistic relativity has been applied before to languages of mathematics, i.e., to formal systems, but such an application is interesting, in part because of evident similarities to conclusions reached by Gödel.

In the sample epistemological analysis that follows, mathematics and linguistic relativity are, by way of conceptual experiment, placed in close association.

Initial Claim: The knowledge a mathematician is capable of attaining in mathematics is limited by the structure of the formal and natural languages which he can use to report his knowledge. I.e., there are certain mathematical truths which a mathematician cannot know because of the limitations of the

languages available to him, yet he can know he is thus limited. (Cf. third example, p. 54.)

I. Identification of what must be presupposed as knowable in order for the initial claim to be made:

a. It is possible to know what the "limitations" are of the languages one uses, of one's total "linguistic répertoire".

b. It is possible to know that there exist mathematical truths not formulable by means of the languges one uses.

c. It is possible to know that the limitations identified under a. are responsible for the impossibility of knowing certain mathematical truths referred to under b.

II. What kind of justification in principle must be possible to know a., b., and c.?

d. To justify that a. is true, it is necessary to be able to know that a. is true and to formulate the claim that a. is true by means of some language that one uses.

e. To justify that b. is true, one must be able to refer to some mathematical truths which one cannot formulate by means of the languages one uses.

f. To justify that c. is true, one must be able to show that the mathematical truths referred to cannot be formulated by means of the languages one uses.

III. Formulating a position:

a. - c. must be presupposed as knowable for the initial claim to be made, and d. - e. state preconditions which must be met to justify a. - c. Do these requirements d. - e. pose impossible demands? If so, then the initial claim is unjustifiable in principle and cannot meaningfully be made. If not, then the initial claim expresses something which can be known about mathematics, and is epistemologically valid.

Decision or metaclaim: d. - f. are demands which <u>cannot</u> in principle be satisfied.

IV. Argument to demonstrate that the above metaclaim compels assent.

Let L be the only language (or the set of all languages) one has available as a frame of reference to express claims. To know that

L is limited in the sense intended is to know that there are mathematical truths which one cannot refer to using L. Let 'M_t' stand for the set of such mathematical truths. 'M_t' is a name belonging to language L. 'M_t' cannot serve both to name mathematical truths which cannot be referred to given the limited resources of L, and yet be a name in L. It must be a name in L in order to conform to the initial stipulation, that L is the only linguistic frame of reference available. Yet it cannot be a name in L if 'M_t' names what the "limited resources" of L cannot name. Hence, d. and e. are unjustifiable in principle because they deny what must be presupposed in order for the presuppositions they are intended to justify (a. and b.) to be sustained. Hence, f. cannot in principle be demonstrated because of the self-referential inconsistency of a. and b. together.

Sketched quasi-formally (additional steps, modal operators and operations are omitted for clarity):

P: Any claim made by George, a mathematician
L.R.: Linguistic relativity is asserted

	Informal statement
1. $P \rightarrow P \varepsilon L$	George's claim entails that his claim is expressed in language L. (Hypothesis, only as informational background to establish the context of reference)
2. $P \rightarrow L.R.$	The linguistic relativity claim is presupposed by George's claim. (Hypothesis, as above)
3. $L.R. \rightarrow -(Q \varepsilon L)$	The linguistic relativity claim entails that reference is made to one or more objects which L cannot make reference to. 'Q' names such an object or objects. (From the analysis of I b., c., and II e., f.)
4. $Q \varepsilon L$	George is forced to use L to express this latter claim: L is his only linguistic resource. (Hypothesis)
5. $L.R. \vdash -(Q \varepsilon L); (Q \varepsilon L)$	The linguistic relativity claim leads to inconsistency. (Elementary logic) Its justification requires use of a referring expression which is ruled out by the hypothesized "limitations" of L.

6.	$-(P \rightarrow L.R.)$	If 2. above is not rejected, any claim made by George will lead to inconsistency. George cannot consistently formulate the linguistic relativity claim: in attempting to state that position, George is compelled to deny that his assertion of linguistic relativity presupposes linguistic relativity. The linguistic relativity hypothesis is therefore self-sabotaging.

Exercizes

Two claims are described below. For each, (1) analyze the claim from an epistemological point of view, identifying what must be presupposed as knowable in order for the claim to be made. (2) Indicate what kind of justification in principle must be possible in order to know these things. (3) Decide whether such justification really is possible. (4) Construct an argument to demonstrate your own metaclaim from (3).

Problem 1. The solution to a set of three simultaneous equations, e.g., is always discovered. It is never invented. This is true whether or not these particular equations have ever been seen by any mathematician before. (Cf. first example, p. 54.)

Problem 2. Let T be a theory which makes it possible to explain an end-state in a problem-solving process, by indicating what earlier steps would in fact lead to the end-state. Given a particular end-state, s, the earlier steps described by T were involved, although perhaps unconsciously, in a human subject's thought process.

Some Epistemologically Problematic Issues

Five different contexts are described below which commonly occasion knowledge claims which in principle cannot be justified. For each of the views described, decide whether you interpret the view in a way which renders justification possible. If you believe that it is possible to justify a particular view, develop a set of reasons why you would claim this. If you believe an epistemologically pathological claim is in question, attempt to justify your position, and then propose any changes in the pathological claim which would, if possible, place the claim on firmer ground.

1. We often claim that certain aspects of a thing or idea are "given implicitly" since, upon analysis, we come to see these aspects, and believe them to have existed prior to noticing them. We thereby come to distinguish between what is explicit and what is implicit. What, in general, justifies the claim that what we perceive at a certain time existed prior to our perception of it?

2. It has become increasingly common to think of our methods of analyzing an issue as "perturbing" the nature of the issue itself. This "perturbation hypothesis" is maintained with respect to physical interactions with subatomic particles when we seek to measure, e.g., their momentum or position. Similarly, we sometimes regard our language as "perturbing" the way we understand reality (the linguistic relativity hypothesis). Or, we may conceive of our theories as determining or interferring with an unobstructed view of reality. How would you analyze this view from an epistemological standpoint?

3. When a phenomenon has been understood in terms of rules which enable us to predict the occurrence of similar phenomena, we are inclined to think that the lawlike regularities we have identified in antecedent phenomena are causal patterns which result in the occurrence of the later phenomena we have predicted. On the one hand, we have a schema of representation, a theory, in terms of which we can make certain predictions, and can confirm or disconfirm these predictions in relation to the occurrence of phenomena we observe. On the other hand, we believe that temporally antecedent phenomena cause the occurrence of temporally subsequent phenomena. Can either or both of these views be justified? How?

4. When we claim to have made a mistake, we usually give priority to a realization we have made later in time. We tend to apply what we <u>have come to see</u> to what we <u>were not able to see</u>. I might say, "I thought that person was Joe, but I was wrong. He wasn't

Joe at all." What justification is there, in principle, for the view that what we perceive earlier in time was in error?

5. In what sense are the claims you have made in connection with 1. - 4. "merely relative" to the frame of reference you have chosen in order to justify these claims? Are your claims "provisional" or "tentative"? Are they "absolute" and "true for all possible worlds"? Are they "controversial" or do they "compel assent"? When you answer these questions, is the justification you offer itself "relative" or "absolute"?

Symptoms of

Epistemological Pathology

Ten groups of concepts and claims are listed below. Some groups overlap. All can, depending upon their employment, involve conceptual pathologies. Each group is preceding by a list of the specific concepts which are embedded in the claims that follow. When, in future readings and conceptual analyses, you encounter these concepts, you should pause, reflect, and seek to determine whether the views you are studying may be self-sabotaging. Frequently, the use of these concepts by an individual theorist serves as a sign or symptom of the presence of conceptual dysfunction.

Study each group of claims. Spend a few moments reflecting on each claim, and attempt to determine precisely why it may bring about different forms of conceptual short-circuiting. The time you spend doing this will be indispensable in your susequent work in epistemology, because you will need to rely on your sensitivity to pathologies of these kinds in order to construct your own epistemological analyses.

1. Concepts: acts, activities, agency, volition, etc.

   a. "When I undertake to remember X's name, a 'faculty of memory' is exercized. It is because of the activities of this faculty which I possess that I am able to remember X's name."

   b. "When I raise my arm intentionally, my will or volition is responsible for the resulting action."

   c. "Man, at least sometimes, is free."

2. Concepts: the subject, ego, person, or agent.

   a. "'I', my 'self' is different from 'my consciousness'. 'I' stand apart from what I am conscious of."

   b. "I am more than a body, a collection of abilities and memories. I am a person, an ego, an agent."

   c. "'If I hit a pane of glass forcefully with a steel rod, it will shatter.' 'If I am hit, insulted, and criticized, I will feel anger, humiliation, or shame.' These two sentences express very different sorts of things: they are not analogous in meaning."

3. Concepts: pertaining to relational ascriptions: "is independent of", "is dependent upon", "is influenced by", "is determined by", or "is caused by".

   a. "She is sad because her mother died."

   b. "Voting should be encouraged in a democracy. An individual's vote affects or influences the kind of social and political environment in which he lives."

   c. "Charles has now become independent of his mother's attitudes; although some of his attitudes resemble attitudes of his mother, she no longer has any power over Charles' outlook."

4. Concepts: "discovery" as opposed to "invention".

   a. "Hemingway did not discover the plot for his novel, For Whom the Bell Tolls, he created it."

   b. "The solution to a set of three simultaneous equations, for example, is always discovered; it is never invented. This is true whether or not these particular equations have ever been seen by any mathematician before."

5. Concepts: relating to genetic claims, e.g., "came to be as a result of".

   a. "Without a great deal of prior work Poincaré would never have been able to generate a proof for such-and-such a theorem. That is, his earlier work served as a necessary preparatory basis for his later insight."

   b. "Cognitive development proceeds through a succession of phases. These phases comprise a process that leads to the improvement of the cognitive abilities of any normal subject."

6. Concepts: linked to explanation, often causal and productive.

   a. "Let T be a theory which makes it possible to explain an end-state in a 'problem-solving process', by indicating what earlier steps would in fact lead to the end-state. Given an end-state s, the earlier steps described by T were involved, although perhaps unconsciously, in a human subject's thought process."

   b. "This volume of water was vaporized because it was heated."

7. Concepts: describing relationships between theory and subject-matter, between mind and physical brain, between rules of usage and ordinary use, etc.

   a. "The properties ascribed to a phenomenon p from the standpoint of a certain theory T are properties which p does not really have."

b. "A man cannot think without a physical brain."

c. "An adequate formulation of the rules of usage for a language L informed us what rules any speaker of L actually follows."

8. Concepts: the unconscious, hidden determinants, etc.

a. "Provided it is possible to explain the occurrence of a dream by making recourse to 'hidden determinants' (a rationale which the subject reporting the dream was not conscious of), then it can be meaningful to claim that the subject had this dream because of these 'hidden determinants'.

b. "The 'spontaneous occurrence of insight' experienced by a mathematician can be explained by postulating unconscious operations which generated the insight. Indeed, all such 'flashes of insight' require unconscious operations on the part of the mathematician."

9. Concepts: suggesting that perceptual structures and organization are "contributed by" or "imposed by" "the subject's activities".

a. "Hilbert's perception of principles of formal organization is to be understood as a contribution to or imposition upon experience resulting from activities of the subject: What Hilbert perceived was not already there, but resulted from a contribution or imposition of order for which he was responsible."

b. "Human experience is spatial and temporal because the human subject contributes or imposes a spatial and temporal structure upon raw data received through the senses."

10. Concepts: used to assert that the results of an analysis, investigation, set of operations, etc., can be given retrospective validity (frequently related to "discovery" claims).

a. "Assume that a computation is made which gives a result R. We double-check R. We prove that -R leads to contradiction. Hence R was true before the calculation was made."

b. "An epistemological analysis of the presuppositional structure of a particular knowledge claim in mathematics leads to a result R. We find that -R leads to self-referential inconsistency. Therefore, R must be true, independently of our epistemological analysis."

Argument Assessment Guidelines

Individual knowledge claims are always stated in relation to some background context of reference, yet rarely will the relationship be made explicit in the form of a structured argument, with a set of premisses clearly identified followed by one or more conclusions. For the purposes of epistemological analysis, it can be helpful to represent a particular knowledge claim in the position it has in an overall argument. The argument must often be reconstructed, sometimes even creatively supplied, in order to represent fully or adequately a particular knowledge claim that is to be analyzed.

The following guidelines assume that an argument can be identified or constructed in relation to which you can situate the knowledge claim you are interested in. The guidelines raise a variety of questions and make alternative suggestions which may be helpful to you in organizing and developing an epistemological analysis.

Understanding the Argument

Step I

First you have to understand the argument

What is the object of the argument? What does it seek to demonstrate? What is the basis for the demonstration? Identify the premisses. Are they compatible with one another? Are any of the premisses themselves inconsistent or subject to controversy?

Separate the various parts of the argument. Are the premisses and conclusions tied together by a relation of logical entailment? Understand the nature of this relation: Is it merely a relation of belief? of fact? of logical necessity?

Devising a Plan

Step II

Analyze the connection between the premisses and the conclusion. You may be obliged to consider auxiliary arguments if an immediate connection cannot be found. You

Once you have understood the connection between the premisses and the conclusion of the argument, you will want to develop a strategy that will make it possible for you to assess the argument critically. Here are some alternative strategies:

(a) Analyze the argument in terms of its

should obtain eventually a plan for an assessment of the argument.

own internal consistency. Do any of the premisses conflict with each other or with the argument's conclusion? Do any of the premisses taken together preclude the success of the argument? Note which basic concepts in the argument remain undefined.

(b) Analyze the argument by relating it to a similar argument the validity of which you can more easily prove or disprove.

(c) Develop an opposing argument that seeks to demonstrate either the opposite of one or more of the premisses that are essential to the argument in question, or the opposite of the conclusion. Show how your own argument relates to the argument you are analyzing.

(d) Develop a defensive argument that seeks to demonstrate that the denial of one or more of the premisses, or that rejecting the conclusion, results in inconsistency.

If you cannot develop what you consider to be a good strategy for assessing the argument, try to remember a similar argument you may have encountered before. Can you imagine a simpler, more accessible, or more specific argument? A more general argument? An analogous argument? Can you devise a plan to assess part of the argument? Try to see how this part which you can handle relates to the overall argument: this may lead you to a more comprehensive stategy.

Are all the premisses required to establish the conclusion? Have you taken into account all essential notions involved in the argument?

Carrying out the Plan

Step III

Construct your own assessment, following your plan.

As you carry out your plan of the argument assessment, check each step. Can you see

clearly that the step is valid? Can you prove that it is valid?

Looking Back

Step IV

Examine the result obtained.	Can you check your result? Can you support your own assessment by making this check explicit?
	Can you derive your result differently? Can you see it at a glance? Can you use the result, or the method you have employed, for some other problem?
	Seek to generalize what you have done.

[The guidelines presented here owe much to George Polya, who developed analogous guidelines for use in mathematics. See George Polya, How to Solve It: A New Aspect of Mathematical Method (Princeton, N.J.: Princeton University Press 1945, 1973).]

Instructions for Argument Assessments

If you are asked to present epistemological analyses before your class, the following instructions may be helpful to you.

An epistemological analysis should attempt to do three things:

(i) Provide a careful but short re-statement of the position of the author in question: Reconstruct in your own words what the author claims, how he justifies or argues his point. Avoid using technical terms; define, in your own words, any technical terms you do use. Remember, the purpose of this part of your analysis is to give the other members of the class a concise exposition of the author's overall position, if possible breaking down the main argument in a step-by-step outline fashion. Spend about one-third of your available time on this part of your epistemological analysis.

(ii) Isolate a specific knowledge claim made by the author and make clear why, according to the author, it should be accepted. Here, you will either be reconstructing the author's own argument, or formulating an argument yourself on behalf of the author's claim. Since you are dealing with a part of the author's overall argument that you already described in part (i), parts (i) and (ii) should dovetail easily.

(iii) Now, analyze the argument you have described. Construct your own assessment to show that the author's claim is or is not validly argued. Formulate your own reasons for claiming it to be valid or invalid. Seek to prove your point, using logical, meaning, or modal criteria.

General Instructions

Use the blackboard (or hand out an outline of your argument). It is important that what you are saying be communicated, so make every attempt to understand your own thinking in clear terms so that others may follow you. Write down the author's main premisses and conclusion as you understand them, and do the same for your own argument.

Avoid subjective reactions to arguments, try to eliminate statements of personal beliefs. Your reasoning should be objectively valid.

On the use of examples: Examples help to illustrate what you are saying, but they cannot by themselves prove a thesis. The same is true in appealing to facts. We are all aware that the certainties of today are often transformed into the dogmatic errors of yesterday. Consider: "the world is flat", "man cannot fly", "no man has lived beyond 150 years of age", "all events are caused", etc. Simply because most people at a particular time in history see things in a certain way does not show their beliefs to be true.

From Dangerous Pitfalls

to Firm Ground

Being epistemologically self-conscious requires a high degree of deliberation and caution when making claims of any kind: Even when you wish to take some claim for granted (as a premiss) in the course of an argument you are developing, it is your responsibility to make explicit (1) what conditions must be satisfied for that claim to be justified, i.e., for you to know what you are talking about, and to show (2) whether in principle these conditions can be satisfied. Unless you do both (1) and (2), your analysis will not be epistemologically self-conscious in the sense we have developed.

The following list summarizes a number of concepts and terms which we saw are frequently epistemologically problematic: This is not to say that such concepts and terms ought to be avoided entirely, but that when you choose to use them, you should exercize caution, and be able to make explicit (1) and (2) above.

Concepts and terms which often are used in conceptually self-sabotaging ways include those which relate to:

1 - acts, activities, agency, process, volition, etc.;

2 - the subject, ego, person, or agent;

3 - relational ascriptions: "is independent of", "is dependent upon", "is influenced by", "is determined by", or "is caused by";

4 - "discovery" as opposed to "invention";

5 - genetic claims (often combining concepts drawn from 1-4);

6 - explanatory claims (often causal, productive, etc.);

7 - relationships claimed between theory and a subject matter, between mind and physical brain, between rules of usage and ordinary use, etc.

8 - "the unconscious", hidden determinants;

9 - the idea that perceptual structures and organization are "contributed" or "imposed by" "the subject's activities"; and

10 - claims which assert that the results of an analysis, investigation, set of operations, etc., can be given retrospective validity (frequently related to "discovery-claims").

This is a list of epistemological "red flags", as it were. When you encounter

concepts used in these ways, a conceptual reflex should tell you to tread with particular care. For at such times, it is especially easy for habitual patterns of thought to undermine intellectual control.

Once one has become sensitive to potential epistemological difficulties, it is heartening to enumerate some of the principal ways in which knowledge claims can successfully and clearly be justified. These modes of justification mark out for us some of the areas of firm ground on which it is possible to take a stand with a more relaxed degree of conceptual vigilance.

There are a number of general and sometimes overlapping areas of epistemologically firm ground; here, we will consider three of these.

1. Functional descriptions

   A knowledge claim can frequently be justified if the use of a particular set of procedures, operations, methods, etc., can be shown to be "functionally basic" to that knowledge claim. For example, knowledge claims which concern alleged changes in physical dimensions are a function, in this sense, of means by which measurements can be made.

   Or, consider my knowledge claim, as I write this: "I am holding a pen." The fact which I assert is ascertained as a function of senses I rely upon, which give me visual, tactile, and kinaesthetic information.

   In elementary arithmetic, to take another example, sums of sets of numbers are determined as a function of the operation of addition.

   As another example, consider that you are making certain observations, and you claim to detect particular changes over time in the phenomena you observe. These changes can, in principle, be noted only as a function of capabilities which permit you to compare temporally successive states. We often do this by means of memory, or we use certain tangible "artifacts", e.g., photographs, which enable us to contrast earlier and later states.

   When a knowledge claim concerns objects of reference which are functionally defined in this sense, the knowledge claim in principle is capable of being justified. In such cases, we know precisely what it is that we are talking about, and how to go about justifying our claim to know this.

2. Dispositional descriptions

   Some knowledge claims involve the ascription of certain properties or relations to an object provided some set

of conditions is met. For example, claims like the following are often made: "I will see more of the desk if I walk around it." "Brittle objects will tend to break when struck hard." "John will get angry if he is teased." "Theorem number 23 can be proved provided one does such-and-such." Etc.

The justification for such dispositional claims is to offer a description of the conditions which would need to be satisfied in order for the expected result to obtain. Such a description is ultimately a description of a set of expectations; if the conditions are satisfied and the expected result occurs, these expectations are fulfilled.

Strictly speaking, it is not possible of course to know that a set of expectations will, in a particular instance, be fulfilled. However, what we usually have in mind is a claim which summarizes, so to speak, our experience with similar phenomena in the past. All objects we have experienced in the past which were brittle, did break when struck hard. This is such an object. I therefore expect it will break if it is struck hard. (See historically relative descriptions, below.)

It is straightforward to describe one's set of expectations concerning a particular object of reference, in the light of one's past experience. Such a description characterizes the basis for a knowledge claim of this variety.

3. Historically relative correlations

Natural science is predominantly concerned with identifying correlation patterns. For example, we claim to know that men die when their hearts stop. Unsupported rocks fall in a gravitational field. No terrestrial bird weighs more than a thousand pounds. --All of these claims express, in lay terms, correlations that have been observed in medicine, physics, and zoology. Of course, a great deal more can be said about these correlations. We could expand our focus to include data relating to oxygen requirements of brain cells, or consider the metric of local space curvature as a theoretically equivalent description of gravitation, or calculate the weight to surface-area prerequisites to achieve flight. Whatever the level of sophistication, the lawlike regularities natural scientists are most interested in constitute patterns of correlation.

From an epistemological point of view, the context of justification for knowledge claims which describe correlations -- men die when their hearts stop -- is history.

> History is the record of our experience. If we restrict knowledge claims that describe correlations to their appropriate context of justification, we restrict the patterns of correlation observed by science, to history. And it would then make sense to consign our scientifically-grounded expectations about the future to psychology. Correlations that have been observed constitute facts about the world. Predictions that assert that such correlations will continue, constitute facts about our minds and our models of reality. Whether the world will continue to bear out what our minds expect is entirely in the world's hands!

In general, the above three modes of justification -- involving descriptions of functionally-based assertions, of dispositional relationships, or of historically relative correlations -- constitute explicit forms of justification. They provide contexts of reference in which a great variety of knowledge claims can be justified, whether they refer, e.g., to controlled observations, everyday sensory data, calculations, general theoretical representations, or declarative assertions in such frameworks as mythology, religion, literature, etc.

Contextually Relative Justification

All of the varieties of justification that we have described above are contextually relative. These modes of justification are available whenever we refuse to allow ourselves to extend the range of application of a knowledge claim beyond the context in terms of which it can be justified. We are, in all three cases 1-3 above, on epistemologically safe ground. For example, the claim, "King Leer was thought to be mad" is a true claim, having Shakespeare's work as its context. Its truth is determined by reference to that context. "Pegasus exists" is a mythological claim, true in that context of reference. "The Tao is spontaneous" is a claim which can be justified in a certin religious or philosophical context. "There is a successor for every positive integer" is a knowledge claim the appropriate context for which is number theory; relative to that context, many mathematicians know how the claim can be justified. It usually takes the form of a functional description. "In the brain, neural excitation is both synaptic and ephaptic" is a claim which has neurophysiology as its context of justification. There, observational data and detection procedures for small-scale electrical impulses constitute the principal means of justification employed.

Any claim to knowledge is associated with an appropriate context of reference. It should make good sense that justification for a particular knowledge claim is, in principle, always relative to that context. It is when we seek to extend our claims to know beyond the contexts in terms of which our claims can be justified, that we become projective in our thinking.

--Projective thought seems to be a natural tendency in man, a kind of epistemological megalomania, a conceptual ambitiousness which is excessive in proportion to what is possible. When our systems of commitment are projective, our attitudes and behavior which express our beliefs then frequently become self-sabotaging. We attempt to do what, in principle, cannot be done: we pull the conceptual basis out from under our own conceptualizations.

To refrain from conceptual immodesty is perhaps the surest path to epistemological sanity.

CHAPTER 7

Postscript

On Framework-Relative Epistemology

Postscript

On Framework-Relative Epistemology

Conceptual therapy is an unusually flexible approach to epistemological analysis because it involves a framework-relative reflection. Whatever the context of reference in view, it is possible to study preconditions which must be satisfied if reference within that context is to be possible. Sometimes, concepts, terms, and claims are used in ways which conflict with the preconditions of their meaning. There are biconditional relationships between the epistemological validity of a knowledge claim, its justifiability in principle, and its meaningfulness.

Conceptual therapy embraces a pluralism of modes of justification. On the one hand, conceptual therapy seeks to identify, reject, avoid, and replace claims for which no justification in principle can be provided. Such claims are conceptually self-sabotaging, and, as we shall see, are meaningful in appearance only. (See Chapter 12.)

Some conceptually pathological claims are relatively harmless: they are without real use or applicability. There is the "epistemological question" par excellence -- Does a tree that falls in a forest where no one is present make any noise? Whether the question is meaningful or not will make very little difference to most people. But many projective claims are among the most destructive, and rarely recognized, forces that exist, forces which bring bitter unhappiness, suffering, and death. Wars are fought in the name of allegiance to absolutist ideologies. The boundaries of their systems of belief are forgotten, and what the pawns of projective commitment do on behalf of a conceptual mistake, entire armies cannot justify, for they know not what they do.

We have acknowledged that there are only circular reasons which rationality can give to persuade the adoption of rationality as a standard for what we think and do. Epistemology is not enough; humanity is essential. Conceptual therapy, as I have attempted to describe it, is a humane and exact discipline which sees the value of linking rationality and commitment, philosophy and the practice of life. In some ways, conceptual therapy is a critical and intolerant enterprise because it rejects self-destructive patterns of thought. In other ways, conceptual therapy is pluralistic, open, and tolerant, accepting the framework-relativity of knowledge claims which reflect a wide range of contexts of reference.

The conceptual double-bind which is, so to speak, inflicted upon individual thinkers by the practice of epistemology, is a benign bind. It captures man in the confines of his own mind to disclose to him the limitations of his own thought. The positions he takes, the claims he makes, may at times compel

assent. Denying them may short-circuit the system of concepts he employs, showing that the fabric of his vision is torn if they are not accepted. But it is important to remember that assent is compelled relative to a framework, a body of theory, a set of presuppositions, an attitude, a model, a subject-matter. Yet this relativity does not make our claims to know "merely relative" in the sense of "tentative" or "provisional". There is no stronger and more convincing demonstration than one which establishes the relativity of a certain variety of knowledge to a chosen frame of reference.

PART II

APPLICATIONS AND EXTENSIONS

We are rapidly reaching a stage when experimenting with theoretical formulations is becoming a distinct discipline in itself.

(William Bender: Introduction to Scale Coordinate Physics)

In the absence of clear knowledge of the meanings and relations of the concepts that we use, we are certain sooner or later to apply them wrongly or to meet with exceptional cases where we are puzzled as to how to apply them at all.

(C.D. Broad: Scientific Thought)

APPLICATIONS AND EXTENSIONS

This part of the book is an attempt to illustrate some of the techniques of epistemological analysis which were described in the first part, and to extend that discussion somewhat more technically. The papers included here will give you some idea of the style and type of work likely to be encountered in professional journals in this field.

The next chapter, 8, presents a general introduction which reflects on the epistemological pluralism that is fundamental to the approach of conceptual therapy.

Chapter 9 formulates the basis for such an approach, describing what I have called a general metalogic of reference.

Chapter 10 views this metaframework in the light of early phenomenology, and applies the resulting approach to a study of the distinction between implicit experience and explicit reflection.

Chapter 11 briefly continues the analogy with early phenomenology, and then turns to a general review of the history of self-referential argumentation. The idea of a metalogic of reference is developed further. The last two sections of the chapter describe applications of self-referential argumentation to several problems in the philosophy of science: to proposed rejections of scientific objectivity, to the doctrine of radical meaning variance, to the Quine-Duhem thesis, and to an analysis of hidden variable theory in quantum mechanics.

Chapter 12 develops a systematic argument to show that referential consistency is a rationally compelling criterion of meaning. An attempt is made to give more formal elaboration of the idea of a metalogic of reference.

CHAPTER 8

Towards a Unified Concept of Reality

TOWARDS A UNIFIED CONCEPT OF REALITY

STEVEN BARTLETT*

It is increasingly less appropriate or useful to speak of a single concept of reality which we generally share. We rather tend to acknowledge the existing heterogeneity of knowledge. Where Aristotle's logic possessed lone authority for some twenty centuries, there is now a growing multitude of distinct systems of logic. Euclid's geometry has become a special case in a family of divergent geometries. Newtonian mechanics remains no more than a reasonable approximation for the purposes of earth-bound engineers, and must be replaced by quantum theory in the domain of the very small, and by relativity theory in the domain of the very large. Etc. This expanding pluralism has been accommodated in a very natural manner by the "systems-approach", but without an attempt to articulate a unified concept of reality.

As a result of the increasing consciousness of pluralism, philosophical obituaries for the doctrine of absolute truth have gradually appeared, albeit prematurely. The doctrine does not give up its ghost easily, in spite of the popular tendency both to emphasize *conventions* and *language-games*, and generally to *relativize* in terms of *conceptual frameworks*. The near-deceased is bitter and ironic: "Without absolute truth, there can be no unified concept of reality."

Of course, the terminally ill egotist often posits his own indispensability — and here the malady is a solipsism in which the doctrine believes itself to be the sole true doctrine about truth. To argue with a dying doctrine is a delicate affair, but the truth is that the doctrine *is* dispensable and, ironically, laying it to rest opens the way to a unified concept of reality.

Perelman has observed that rhetoric has been criticized by those "for whom there was but a single truth in every matter." [p. 45][1] This seems to suggest that rhetoric may sanction more than one truth in a single matter. Certainly this view seems to follow once one admits the existence of a plurality of sets of rational first prin-

*Visiting Research Fellow, Max-Planck-Institut zur Erforschung der Lebensbedingunen der wissenschaftlich-technischen Welt, Starnberg, West Germany.

ciples, where each set provides a framework for an appropriate set of truths relative to it. Perelman's familiarity with formal systems and with variations in judgments from epoch to epoch, and from culture to culture, doubtless provides him with evidence for the existence of such a plurality. Furthermore, it is basic to his concept of rhetoric to "combat uncompromising and irreducible philosophical oppositions presented by all kinds of absolutism . . .". [p. 150][2]

Perelman proposes the notion of the audience-relativity of meanings and usages. (For his remarks on the audience-relativity of rhetoric, see [pp. 121, 138][1] and [pp. 7, 14, 19, 21, 39, 54, 65ff, 72, 110, 134, 507, and passim.][2].) His thesis immediately gives rise to the question whether there is but a single truth for all audiences, or whether there are various truths, each relative to some ideal group of similarly constituted individuals.

Although Perelman apparently wishes to avoid any form of absolutism, he does speak of "the universal audience" in terms of "that invariant faculty, present in every normally constituted human being, which is reason." [p. 127][1] (Of course, what must be qualified as *normally* invariant is an odd "invariant" indeed!)

Perelman remarks:

> When a stick is partly immersed in water, it seems curved when one looks at it and straight when one touches it, but in reality it cannot be both curved and straight. While appearances can be opposed to each other, reality is coherent: the effect of determining reality is to dissociate those appearances that are deceptive from those that correspond to reality . . . (B)ecause of their incompatibility, appearances cannot be accepted together . . . (p. 416)[2]

This concept of reason functions as a norm to eliminate or smooth out incompatibilities in "appearance" by distinguishing data to be retained as significant from data to be rejected as misleading. Such a norm opposes, in an absolutist fashion, the simultaneous truth of both terms of an opposition. The stick "cannot be both curved and straight."

Nevertheless, *in fact* what we *see* is curved, and *in fact* what we *touch* is straight, yet also *in fact,* Perelman says, the stick *really* is not bent at all. So we soon encounter the need for a good understanding of what a fact *is*. . . .

Before any thesis can be argued, some set of criteria of sound argumentation and some set of facts judged to be relevant must be accepted explicitly or implicitly in advance. Indeed, it seems to be fairly clear that in the selection of "relevant facts" certain of the

basic postulates of reasoning are already involved. In this sense, a fact bears witness to some fundamental postulates a function of which it is recognized to be a fact. In other words, a proposition can be claimed to assert a fact only in relation to some set of norms or standards which are ordinarily agreed upon within a given framework of reasoning. It follows that it must be in relation to such norms or standards that the concept of fact is best described.

We say that what a true proposition asserts is a fact. A proposition which is confirmed, provisionally or otherwise, is asserted to be true. (On presumed versus observed facts, see [p. 74][2].) Thus, "any truth enunciates a fact." [p. 69][3] A fact is not only what a true proposition asserts but it is also what a false proposition denies. Facts are themselves neither true nor false, but they render assertions true or false. (For a related treatment of the concept of fact, see [pp. 177-281][3] and [pp. 85ff][4].)

What makes a proposition true or false is expressed *via* assertions, denials, beliefs, etc. Thus, what makes it possible to identify a fact is a framework in terms of which relations can be established between an individual (who asserts, denies, . . . , a given proposition) and the world of things with which the individual is acquainted. In short, a pragmatical framework (one which, strictly speaking, coordinates persons, meanings, and events) is necessary in order that candidates for facts may be specified and become subject to methods of confirmation.

A framework adopted for the identification of facts defines the factually real in terms of the criteria which it may presuppose for discriminating between significant and misleading data. When we claim that any assessment of the truth of an assertion must take account of the context in which it is made, we adopt a point of view which is "framework-sympathetic". We realize that an examination of the context in which a proposition is asserted will frequently reveal the framework and presupposed standards for the recognition of facts, in relation to which coordinations between persons, intended meanings, and identifiable events are to be understood. A concrete description of such a pragmatical framework determines the ideal audience, to recall Perelman's language, relative to which a set of propositions is acceptable as true. It is an easy step, then, to extend Perelman's notion of audience-relativity and audience-pluralism, to a recognition of the relativity of facts to the ideal representation of a particular framework by an audience.

It is an immediate consequence of this framework-sympathy

that what is factual in relation to one framework may not be factual in relation to another. It is in this sense that relative to one framework for the identification of facts, the visual one, the stick is bent, whereas relative to a second framework, the tactile one, the stick is straight. (A similar example may be had in the contrasting facts ascertained about the nature of light by means of two different experimental apparatuses: one indicates light to be a particle-phenomenon, the other that light is a wave-phenomenon.)

It is tempting to ask how this can be when only one and the same stick (or physical phenomenon) is in question. But such a question is of the "How often do you beat your wife?"-type:it is excessively free in its presuppositions. It supposes that the two sets of facts refer, in fact, to the same thing. Sameness has always been philosophically slippery, and common sense yearns for *the same stick*. The result has been to smooth out the bothersome heterogeneity by an appeal to a higher-order framework capable of absorbing the different facts (I limit myself to "facts" since we are now beyond "appearances"). This leads to the tolerant synthesis: *real stick* — which can be both seen ("as if" bent when partly underwater, "as if" (?) straight when wholly in the air) and touched (normally straight). This move in itself is unobjectionable: we have merely developed a less simple understanding of sticks, water, and air, and feel assured in the precise vocabulary of refraction. But the move to a higher-order framework is accompanied by an exhibition of ontological snobbery: After all, the genuinely real stick concerns us, and not the misleading visual image. (The near-deceased raises himself on an elbow and smiles.)

But we are too hasty. The absolutist here neglects his facts: The stick which is visually perceived is *really* seen to be curved. That is a fact. In the interests of a unitary conception of reality, we have permitted ourselves to reject as misleading any factual variation — aberration — from a *norm*. Refraction theory provides a framework in which a set of facts may be interpreted. What is at issue is a *theory* — a system of interpretation — and not a dogma of revelation.

It is worthwhile noting how a theory succeeds in speaking of "the same thing" — be it a stick or a photon — from different points of view. A stick or quantum event is identified as a function of the operations employed to study it. Refraction provides such an identification procedure by describing, for example, how what we see is a function of the medium through which light passes.

The theory of refraction makes it possible to coordinate the bent stick that is seen with the straight stick that is touched, and to consider both to refer to the same object functionally defined by the theory.

From the standpoint of a theory adhering to the principle of non-contradiction, a proposition and its negation cannot both be true. In this sense, only one of the two propositions, 'the stick which is visually perceived is curved' and 'the stick which is visually perceived is not curved' may be considered to be true.

However, a proposition and its negation may both be confirmed in certain theories. But, as is always the case, attention must be paid to the contexts in which the propositions are ascertained. If each assertion is true relative to a different context of reference, then the facts asserted by the two propositions are called *complementary facts*. Facts asserted by contradictory propositions are complementary provided that each assertion is true relative to a distinct context of reference. Consequently, contradictory propositions which have been confirmed in relation to different modes of observation can be regarded as asserting a complementarity of the facts they refer to. (In relation to distinct experimental contexts of reference, the complementarity of predicates ascribable to light — 'is corpuscular'/'is not corpuscular' — has been observed.)

If it can be granted that there are numerous, distinct systems equipped to ascertain facts, formulate true propositions expressing these, and hence reach "objectively valid results", then we must also accept the fact this view brings to our attention: that there is a plurality of sometimes divergent facts, and that the relations between certain of these facts will be relations of complementarity.

Perelman has argued that "revision [of an axiom] cannot be effected by an argument developed within the system to which the axiom belongs." [p. 105][2] Since Perelman is ambiguous in his apparent acceptance of the doctrine of absolute truth, it is difficult to be sure to what extent adherence to a view of the complementarity of facts diverges from the concept of rhetoric he proposes. It seems likely that some divergence is implied by Perelman's reference to "the problem that is raised by the incompatibility of appearances." [p. 419][2] Several questions come to mind: Is there *in fact* an incompatibility of appearances? If there is, what is the nature of the incompatibility? Could it become problematic? If indeed it can, would a concept of reality which bases itself upon the doctrine of absolute truth resolve the problem satisfactorily? The

foregoing discussion should be able to throw some light on these questions.

I have suggested that when we speak of *appearances* as opposed to *reality* we frequently and mistakenly believe the reports of our senses to be in conflict and to require the kind of smoothing out which the application of a reality-standard is intended to provide. *In fact*, these reports are usually not in "conflict"; often, they yield a recognition of complementary facts. (We should need to stray far from "normal psychology" to find an example of this kind of conflict. I should have, e.g., to see the stick as *at once* bent and not bent. For the sake of simplicity, I have omitted any discussion of this dimension of time: If I see the stick now as bent and later in time as not bent, there is "conflict" between the two observations only provided absolutist reality-norms intervene. These would suggest that the two visual perceptions are of one and the same thing — that if the conditions of observations have not changed, then I am likely to have erred in judging that the stick I first saw was *actually* bent, etc.)

If we are persuaded to accept this view of complementarity, it is because we see that propositions asserting facts become incompatible only if they are ascertained from the standpoint of the *same* frame of reference, only if they are assessed in terms of the *same* standards of confirmation. If the propositions in question occur in essentially dissimilar contexts of reference, it will be illegitimate to place them on the same footing and to judge them with the same criteria. Thus, the question whether facts are incompatible or complementary will turn on whether the facts are ascertained in the same context, or in essentially different contexts.

It follows, then, that incompatibility between *facts* becomes problematic either when they are illegitimately treated by neglecting the dissimilarity of the contexts of reference relative to which they were ascertained, or when they are found to conflict in the same context. And, as I have tried to suggest, a doctrine of absolute truth is entirely unsatisfactory once a plurality of essentially dissimilar frameworks of reference is admitted.

It is therefore my contention that (1) there is frequently no need to discriminate against certain facts because of apparent incompatibility, (2) to do so is often illegitimate since (3) there is in fact a plurality of what I have, with deliberate vagueness, variously called rational first principles, criteria of sound argumentation, basic postulates of reasoning, or reality-criteria.

It is fortunate that theoretical developments have brought with them an unsettling pluralism. As variety is compounded, the desire to smooth out or to reject factual heterogeneity show itself to be stubbornly and blindly dogmatic. There are many points of view, and many facts, perhaps not all of which can be accommodated within the compass of a single theoretical interpretation. But this does not imply that a unified concept of reality is impossible or unrealistic. We can dispense with the idea that unification is possible only at the expense of difference, and consider the sense in which complementarity can accommodate heterogeneity on the terms proferred by this new form of synthetic understanding.

REFERENCES

1. Perelman, Ch. *The Idea of Justice and the Problem of Argument*. Trans. by John Petrie. Routledge and Kegan Paul: London, 1963.
2. Perelman, Ch. and L. Olbrechts-Tyteca. *The New Rhetoric: A Treatise on Argumentation*. Trans. by John Wilkinson and Purcell Weaver. University of Notre Dame Press, 1969.
3. Russell, Bertrand. The Philosophy of Logical Atomism. Reprinted in *Logic and Knowledge*. Ed. R. C. Marsh. Macmillan: London, 1956.
4. Weitz, Morris. The Unity of Russell's Philosophy. In *The Philosophy of Bertrand Russell*. Ed. Paul Arthur Schilpp. Harper and Row: New York, 1963.

CHAPTER 9

The Idea of a Metalogic of Reference

THE IDEA OF A METALOGIC OF REFERENCE *

by

STEVEN BARTLETT
Department of Philosophy, Saint Louis University, Saint Louis, Missouri

Introduction

I would like to address the interests of an approach in philosophy which seeks to disclose and to investigate basic commitments involved when questions are raised about the possibility of experience, the possibility of knowledge, or the possibility of theory in general. A concern for the structure of the possible has, since Kant, traditionally gone by the name 'transcendental'. The basic commitments or investments involved in doing transcendental philosophy will be central to what I wish to treat here. In a sense, then, the context for what follows intends to offer a basis for a metacritique of transcendental philosophy.

I have been persuaded that a transcendental approach can gain a helpful measure of clarity and precision by shifting from the traditional Kantian perspective to a point of view that emphasizes the nature of referring. This shift, as I propose to describe it, provides an effective means for confirming transcendental results. A need for ways to demonstrate the validity of transcendental claims will bring me to a discussion of what I term 'self-validating logics'.

To be specific, I will (i) suggest a rationale behind shifting to the perspective of referring, (ii) propose a general *metalogic of reference* that retains the interests of transcendental philosophy, (iii) describe the usefulness of self-validating logics in this context, and (iv) conclude with some remarks about the value of transcendental philosophy, referring, and the idea of self-validating logics for philosophy of science.

* Research reported here was partially supported by a grant from the Max-Planck-Gesellschaft.

The transcendental approach

Transcendental philosophy finds its orientation in a movement away from a reflection on the actual as such, to a study of the preconditions of its *possibility*. The concept of possibility is fundamental to the transcendental approach, although exactly what possibility *is* has remained vague in the literature of the transcendental perspective. There are a number of alternative conceptions of possibility. I will suggest six of these, in an approximate order of concepts of increasing generality. This sequence will serve to determine a highly general, comprehensive sense of possibility, in terms of which a rationale for the shift I propose will be evident. The alternative views are these:

1. What is possible refers to future alternative states of a physical system.
2. The Stoic-Diodorean view: What is possible refers to what is or will be.
   The Aristotelian-Megarian view: What is possible includes what is, will be, or has been.
3. What is possible relates to the status of a description of an event which is *not excluded* by the known laws of nature.
4. What is possible is classically free of contradiction.
5. What is possible includes those real or abstract objects of reference, of which we can predicate what are ordinarily considered to be incompatible propositions.
   (I have in mind such ascriptions of properties as are frequently termed 'complementary' in elementary particle physics).

To these five views of possibility, a sixth is added that offers some promise as a highly inclusive concept of possibility.

6. What is possible can be understood as a function of an analysis of preconditions of *valid referring*. This view will be developed in what follows.

The general idea of a metalogic of reference

It is convenient to talk about referring in the context of an analysis of descriptions. Both in ordinary usage and in the natural and behavioral

sciences descriptions are relied upon to serve a variety of referential functions.

The referential character of descriptions can be analyzed in terms of the commitments descriptions entail. A description presupposes certain commitments to a framework or family of similarly constituted frameworks. These commitments can be made explicit by thinking, for example, of the general, frequently quite vague, rules or conventions which lend some form of organization to admissible descriptions that can be articulated in the context of a given framework of reference. Since the relationship between conditions of reference and any description is logically prior (in the sense intended by transcendental philosophy in the Kantian tradition) to the formulation of any specific description as a necessary presupposition of it, it seems justified to speak of "referential preconditions".

Referential preconditions are restrictions The hierarchy of different concepts of possibility 1. - 6. is actually a list of various *ways of enforcing restrictions* as to what sorts of possibles we are prepared to speak of. So, an interest in preconditions of reference can be understood as an interest in sketching out a certain sort of general map of a domain of objects for which we want to assure the possibility of valid referring.

These preconditions of reference can be approached in either of two different ways: On the one hand, a study may be undertaken of a specific framework of identification: e.g., the framework presupposed in developing a general phenomenology of human visual perception, or the framework presupposed by quantum mechanical descriptions, involving the use of special kinds of measuring devices as well as an explicit or implicit theory of measurement which permits the significant use of apparatus and interpretation of observations. On the other hand, a study may be undertaken of the very general principles which seem to underlie an entire group of special identification frameworks: e.g., from the standpoint of a phenomenological account of objectivity, the group of identification frameworks - visual, auditory, tactile, etc. - that together provide a basis for the constitution of objectivity, or alternatively, the family of conceptual frameworks with which we are acquainted in the natural sciences, which together determine what is to be understood by 'nature'.

It is in this second sense - the sense in which a study is possible of the general principles of reference that underlie a group of identification

frameworks - that it is appropriate to speak of a general *metalogic of reference*. At this point, then, a metalogic of reference appears to furnish a context for a reflection on the nature of theories in general, where specific cases may be a theory of experience, a theory of knowledge, or any of the various natural or behavioral scientific theories.

Initially, then, my interest is a purely abstract one - without regard for any special theoretical identification framework; without attending, at least in the beginning, to framework-specific rules and conventions - in short, to study pervasive constraints that condition valid referring.

One approach to these highly general and abstract metalogical preconditions of referring is suggested if we think in terms of the kinds of second-order constraints which first-order constraints of a special identification framework must obey to avoid *self-referential inconsistency*.

What I mean by 'self-referential inconsistency' would involve a more technical discussion than I can undertake here, but the basic idea is simple. It is this: Paul Lorenzen, in a different context, refers to what he calls "elementary ways of speaking". He says:

> the decision to accept elementary ways of speaking is not a matter of argument. It does not make sense to ask for an 'explanation', or to ask for a 'reason'. For to ask for such things demands a much more complicated use of language than the use of elementary sentences itself. If you ask such questions, in other words, you have already accepted at least the use of elementary sentences.*

A self-referentially inconsistent use of elementary sentences in Lorenzen's context would involve the decision to employ elementary sentences in doubting the justification of using them.

The main difference between Lorenzen's view and the idea of a metalogic of reference lies in this fact: In a metalogic of reference we are concerned not with elementary usages of the language we, in fact, employ, but with "elementary" means of referring of such a kind that they immediately are involved if we consider referring as a pure possibility. In other words, the very *possibility* of calling such means of referring

* Paul Lorenzen: *Normative Logic and Ethics* (Mannheim/Zürich: Bibliographisches Institut 1969), p. 14. Cf. also P. Lorenzen: *Einführung in die operative Logik und Mathematik* (Berlin: Springer 1969).

into question presupposes them as elementary.

It is here that the idea of a metalogic of reference can be developed by resorting to what I call *self-validating logics*. I am motivated to talk about *logics* in order to furnish a context-relative means to test the same kinds of claims which a Kantian transcendental deduction seeks to justify. A self-validating logic, unlike a transcendental deduction, is fairly simple.

To make clear what I have in mind, let us suppose we wish to study what we believe to be a basic premiss of referring:

If we assume we want to think or talk about a collection of objects of various sorts, we are compelled to allow some means for this thinking or talking about them to proceed - we must be permitted somehow to refer to what we want to think or talk about. This is trivially true, and therefore I take it as basic.

Consider a candidate for a postulate in a metalogic of reference: If a metalogic of reference is to constitute a self-validating logic (or family of logics), then its axioms and postulates will themselves be self-validating, in this sense:

> *A postulate is self-validating if its denial will result in self-referential inconsistency.*

Let us consider the following as a potential elementary postulate for referring, which it seems apt to describe as a "rule of referential counter-exemplification":

> The assertion of the impossibility of referring to an individual something metalogically implies that reference is made to that thing.

This postulate *self-validates* as follows: Reference must be made to that individual something if it is to be possible to say that reference to it is impossible. The self-validation consists in the fact that a denial of the possibility of referring to an individual something is self-referentially inconsistent.

Now, if for the purposes of my informal treatment here it can be allowed that it may be possible to determine a significant number of self-validating axioms and postulates, and then to relate them in a unified and well-ordered formal system, then we would arrive at the idea of a

self-validating logic. It would differ from an ordinary formal system in that its elementary propositions are not merely postulated with some element of arbitrariness, but present themselves as compelling our assent to them if we are to be *able* to refer at all, somewhat in the manner of Lorenzen's elementary ways of speaking.

Were this to be accomplished, we would gain a significant measure of metalogical understanding of the most fundamental commitments involved in referring, an understanding that can be justified by appeal to self-validating demonstrations.

We are then not very far from being able to apply these results so as to better our understanding of, for example, the fundamental structure of a natural scientific theory. For a theory, there will be some domain(s) of objects in which its interests lie, and there will of necessity be an assortment of ways at the disposal of the scientist to refer to the objects he studies. The scientist is particularly desirous, one might add, of supplying a basis for the kind of referring his formal theory, schema of interpretation, and domain of objects oblige him to have. For, as a scientist, he chooses to respond to a need to bring the referring descriptions for which his conceptual framework provides a basis as close to the ideal of unambiguous identification as possible. And this objective is satisfiable only if fundamental commitments involved in the scientist's use of referring descriptions are made explicit and can be seen not to conflict with the theoretic claims he wishes to make.

Transcendental philosophy of science taken in this sense has several functions: to elucidate the referential preconditions basic to specific theories and shared by groups of theories, to detect self-referentially inconsistent patterns of referring, and finally to suggest valid ways of referring to replace unsound ones. To these descriptive, critical, and prescriptive functions may be added a fourth - a preventative task: to furnish guidelines in the form of a usable metalogic which can serve the interests of self-conscious and consistent theory construction.

The association with the medical model is obvious. The descriptive, critical-diagnostic, prescriptive, and preventative functions in a transcendental philosophical context are intended to contribute to the needs of theoretical soundness; the physician accepts identical functions in attending to the needs of human physical and emotional health. The

analogy between sickness and theoretical inconsistency, between medicine and philosophical therapy may to some comprise a repugnant model, but, at the same time, to deny description criticism, or criticism positive prescription, or positive prescription preventative recommendation, will to many seem arbitrary and irresponsible.

An interesting and useful philosophical reflection on the foundational structure of scientific theories I believe is offered by a metalogic of reference: Critical and close attention would be paid to the interconnection between the ways of referring essential to a theory and the objects to which the possibility of access is thereby assured, and between these ways of referring to a domain of objects, and the interpretation placed upon findings from that perspective. The understanding acquired could not wish to take the place of the natural scientist's own comprehension of his field, but it would be a qualitatively different kind of understanding, perhaps more analytically self-conscious, and ought, one would think, serve to enhance, to complement, and to render more precise the outlook of both unphilosophical scientists and non-scientifically oriented philosophers.

Summary

The author shifts the perspective of transcendental philosophy from its traditional Kantian orientation to the point of view afforded by an analysis of preconditions of referring. This shift in perspective is proposed in order to gain clarity and precision, and to provide a means for demonstrating certain of the results of transcendental philosophy.

An attempt is made to achieve systematic clarity for a concept central to the transcendental approach, the concept of possibility. The idea of a general *metalogic of reference* is proposed as supplying a highly inclusive framework from the standpoint of which preconditions of possible reference can be investigated.

The usefulness of *self-validating logics* for transcendental philosophy is suggested as furnishing a metalogical resource for transcendental demonstration.

The author concludes with a discussion of the value of a transcendental metalogic of reference for philosophy of science.

CHAPTER 10

Phenomenology of the Implicit

Phenomenology of the Implicit

By Steven BARTLETT

Summary

An attempt is made to suggest an alternative approach to certain of the problems central to transcendental philosophy. In particular, it seems that the present understanding of pre-reflective awareness, of reflection, and of their interrelation can acquire a greater degree of rigour and clarity. To this end, attention is paid to *preconditions of reference* that are entailed by the phenomenological distinction between implicit experience and explicit reflection.

Résumé

L'auteur propose un accès alternatif à certains problèmes centraux de la philosophie transcendentale. En particulier, il paraît que la compréhension des idées de conscience préréflexive, de réflexion, et de leurs relations mutuelles, peut être précisée et élucidée. A cet effet, l'attention est attirée sur les *préconditions de référence* qui sont impliquées par la distinction phénoménologique entre l'expérience implicite et la réflexion explicite.

Zusammenfassung

Der Autor versucht, eine alternative Untersuchungsmethode zu bestimmten Hauptproblemen der transzendentalen Philosophie darzustellen. Insbesondere könnte das gegenwärtige Verständnis der vorreflexiven Bewusstheit, der Reflexion, und ihres Zusammenhangs strenger aufgeklärt werden. Zu diesem Zweck wird Aufmerksamkeit auf die *Beziehungsvorbedingungen* gerichtet, die in der phänomenologischen Unterscheidung zwischen impliziter Erfahrung und expliziter Reflexion miteinbegriffen sind.

The task of reflective philosophy, and of phenomenology, in particular, has been variously described, but in most of these accounts, and perhaps in all of them, use is made of a notion that has remained both central and vague. I shall call this the notion of the *implicit*: the relation between what is implicit and what is explicit bears certain important similarities, as I will attempt to show, to the relation between what has been termed " pre-reflective experience " and " reflective experience ".

This paper attempts (1) to describe a framework in which the notion of the implicit can be investigated, (2) to reach certain conclusions about this notion, and, in the course of doing (1) and (2), (3) to throw some light on

how reflective philosophy, and particularly phenomenology, can be fruitfully developed along somewhat different lines than tradition has so far made possible.

Accordingly, my approach will involve three steps: In the first, I will try concisely to enumerate a group of defining properties which can serve to characterize the general framework proper to Husserl's phenomenological methodology. In the second, I will suggest a " transformation schema " that permits, with a consequent gain in rigour and clarity, a transposition or reorientation of phenomenological methodology in terms of a different, though related, framework. In the third, I will suggest certain conclusions about the peculiar notion of the implicit.

It may be of interest to digress at this point before continuing, to remark that one of the implications of the subsequent discussion concerns certain conflicts between the Husserlian and the Heideggerian approaches to phenomenology. Although this question could not be handled in detail here for reasons of space, a few general comments can serve to place the issue in the present context.

It is basic to Heidegger's philosophical perspective (a) to assert that theoretical analyses *per se* fail to disclose those structures which are *most* fundamental in a description of the constitution of experience, while he claims (b) that reflective theoretical investigations are themselves to be understood as in a sense disguised and disfigured representations of the truth about experience.

Husserl's reaction to these dogmas of *Sein und Zeit* is recorded in the so far neglected marginal comments written by Husserl in his own copy of Heidegger's book [1]. Husserl's reaction is frequently not sympathetic: his comments reveal his skepticism that any philosophically meaningful description can be made without recourse to reflection.

If this is so, then the above claims (a) and (b) lead to a curious paradox which can be formulated as follows: For any reflective philosophical description, by (a) there remains something *un-said* which is true, while by (b) what *is* said can have no real, i. e., " authentic ", claim to truth. But the foregoing statement is, by its own admission, unconvincing as it stands.

The paradox may safely be dismissed if it can be shown that it is indeed the case that a reflective standpoint is the pre-condition for the possibility of philosophical descriptions in general. This is precisely the conclusion reached below, in Part III.

[1] These comments have been brought together by Stephan Strasser, and exist as yet only in manuscript form in the Husserl Archives, Louvain, Belgium.

I

The following description of Husserl's concept of phenomenology will be given in terms of a series of selective idealizations of what I believe to be are defining properties of Husserl's position. I use the expression 'selective idealization' in the sense that the group of defining properties enumerated here does not consistently accomodate all of the multiple descriptions chosen by Husserl to represent his approach to phenomenology. This group contains five such defining properties, which will be discussed under the following headings: the level of maximum theoretic generality, the empirical basis, the transcendental, the conversion of meaning, and phenomenology as a descriptive science.

A. *The level of maximum theoretic generality.* The expression used for this heading is borrowed from F. B. Fitch [2], who employs it to refer to the occurrence of a theory in its own subject-matter. Such a theory concerning theories in general is said to be expressed on the " level of maximum theoretic generality ". Such a theory of theory or science of science is self-referent insofar as it is part of its own subject-matter [3].

Husserl advanced the idea of phenomenology as a science which sets the task for itself to study the general nature of all science, and, in particular, to do this by an investigation of the transcendental foundations of, e. g., the various scientific disciplines, its own transcendental foundation included [4].

Thus, Husserlian phenomenology may be characterized initially as a transcendental science of maximum theoretic generality.

B. *The empirical basis.* The range of objects and structures treated in phenomenology is intimately associated with the " empirical basis " provided by the world of facts and the world of fancy, which furnish material for study. Access to individual instances which may be variously observed or examined is presupposed. The concept of " the given " in Husserl's phenomenology will be interpreted as asserting such an access [5].

In its transcendental capacity, phenomenology is specifically oriented toward an examination of the formal constitutive structures of any particular phenomenon in terms of a range of material possibilities. This may be taken in somewhat the sense that " abstract ontological conditions ... refer to concrete ontological situations and cannot substitute for them " [6].

[2] F. B. Fitch, *Symbolic Logic* (New York: The Ronald Press 1952), p. 223.

[3] An examination of problems of self-reference in such a theory is set aside for treatment elsewhere.

[4] Cf. Husserl's *Logische Untersuchungen* I § 42; II. 1 Appendix 2 § 6; *Formale und Transzendentale Logik* §§ 10, 101; *Ideen* I §§ 62, 65; *Méd. Cartésiennes* 130.

[5] Cf. *Log. Unt.* III § 16; *Formale und Transzendentale Logik* §§ 58, 89, 98.

[6] James K. Feibleman, *Ontology* [New York: Greenwood Press 1968; first printed (Baltimore: Johns Hopkins Press 1951)], p. 140.

C. *The transcendental.* Husserl's phenomenology is concerned with the transcendental in the traditional sense that conveys commitment to the work of elucidating conditions of possibility. The task of transcendental elucidation may be thought of as the isolation and descriptive characterization of the structures or properties necessarily invariant with respect to a certain class of objects. In this sense, these structures and properties are thought of as expressing the essential constitution of the class of objects. Membership in the class entails a certain minimal satisfaction of a set of requirements. Such requirements, in other words, condition possible membership in the class. Discrete classes of objects so studied are specifically determined in relation to distinct and invariant features disclosed by given individual instances which may be diversely observed or examined, as already mentioned.

D. *The conversion of meaning.* Phenomenology, as Husserl observed [7], can be understood in terms of a " conversion of meaning " which the assumption of the " phenomenological attitude " determines. Descriptions of phenomenology in these terms have led to unfortunate interpretations of Husserl's phenomenology. The formulation of the phenomenological approach in terms of the *performance* of various *reductions* has been misleading [8].

What is misleading about such formulations is that they have suggested that access to the framework proper to phenomenology can be achieved by means of certain psychological transformations in attitude. This suggestion reflects the " natural standpoint ", whereby naïve and unquestioning use is made of unclear or biased concepts — here, for example, the concepts involved in " achieving a result, performing an operation, and thus changing an attitude " — concepts to be investigated, rather than to be assumed in the investigation itself. Any formulation of the phenomenological approach which is forced to assume and employ concepts which have not been adequately clarified will be of no value.

The " conversion of meaning " which the " reductions " " lead to " must therefore be provided with some admissible interpretation. My suggestion is that the " reductions " be considered to *define* or *determine* a standpoint, rather than to provide access to that standpoint " as a result of the performance of a psychological readjustment in orientation ". In the sense proposed, then, the meaning of ' phenomenology ' is defined in relation to the following schema.

[7] Cf., e. g. *Ideen* I, Einleitung.

[8] Eugen Fink, in his well-known article, "Die phänomenologische Philosophie Edmund Husserls in der gegenwärtigen Kritik", *Kantstudien* XXXVIII (1933), pp. 319—383, has described certain of these misled interpretations.

(i) With a given object, an ideal and perspicuously defined possibility is correlated, as exemplifying an instance of a class of observable (examinable) objects of similar structure. In this, essentially non-observable, or absurd, possibilities are excluded.

(ii) The possibility described is ranked hierarchically in terms of classes of greater and lesser extension. This provides a system by means of which individual phenomenological descriptions can be classified.

(iii) Individual members of the class in question (or " region ", in Husserl's terms) are considered to be determinable in terms of general concepts and principles expressing the invariant structure(s) of those members of the class with respect to that class.

(iv) The essential structure of an object is determined through the elucidation of the essential connection evidenced between the object — as structurally described in (iii) — and its structure given in a particular mode of observation, e. g., perception. In Husserl, " intuition " is frequently the apprehension that this connection obtains.

(v) The essential connection elucidated in (iv) is studied in either one or both of two different ways: (a) in terms of *static constitution,* relating to the conditions of possibility of an object, the structure of which is *given;* (b) in terms of *genetic constitution,* relating to the conditions of possibility of an object, the structure of which reveals a *synthetic-productive activity* of consciousness. For Husserl, an analysis of genetic constitution is often thought to provide a deeper insight into the origin of what is statically described [9].

E. *Phenomenology as descriptive science.* The framework proper to phenomenology is such as to allow for the description of the constitutive structural principles which condition the possibility of phenomena, as essentially examinable objects. The distinctive character of phenomenology is its task of foundational elucidation. Whatever method is proposed to this end must satisfy some standard(s) with respect to which phenomenological descriptions are protected against the introduction into descriptions of pre-analytically accepted positive content. It is in this connection that Husserl characterized phenomenology as " presuppositionless ".

[9] Due to a fundamental inconsistency in the idea of genetic constitution, which I have treated at length elsewhere *[Théorie de la relativité de la constitution phénoménologique* (doctoral thesis, Université de Paris 1970)], (b) above will not be described here as representing an essential charasteristic of the phenomenological approach.

II

The lapses from rigour and from clarity of exposition that are frequently associated with frameworks expressed in the " language of consciousness " are largely to blame for the neglect of Husserlian phenomenology in the Anglo-American world.

In part to attempt to remedy this defect, I will propose at this point a transformation schema in terms of which a somewhat more exact and clearer description of the framework proper to phenomenology can be given in the " language of reference ". This " transformation schema " will be made clear by associating with each of headings A-E a second heading A′-E′.

A′. The initial program of phenomenology, as a science of maximum theoretic generality, is to secure a logically sound methodology which can provide the basis for a meta-science capable of investigating, on a cross-disciplinary basis, concepts basic to, e. g., disciplines in which scientific methods are fundamental. Such a methodology would provide a useful and meaningful theoretical framework and method for the elucidation of the essential constitution of possible objects of reference. The structure of such a reflexive system is of a totally intrinsic kind — that is, self-reference in the system will require recourse to no higher-order referential embedment-system(s) [10]. This will guarantee that such a discipline will be able to investigate the constitutive foundations of the various particular sciences, its own foundation included. In this context, phenomenology re-appears as a transcendental science of maximum theoretic generality.

The world of ideally possible objects of reference comprises the subject-matter for study. Phenomenology elucidates the sense the world has relative to a given theoretical framework. In this, as will be seen, it denies that meaning can attach to any question concerning this world taken apart from an appropriate possible frame of reference.

B′. Analyses of constitutive structures are essentially relative to one or more given systems. The constitution of a specific object of possible reference, relative to a certain frame of reference, cannot often simply be " read off ", for a good deal of analysis is usually first necessary. In this, the subject-matter for analysis remains presupposed as object of possible reference.

For the present, let the term ' phenomenon ' be synonymous with ' possible object of reference '. Now, the formal structure of a group of phenomena differs from the structure of a given phenomenon only in degree of specificity, that is, in the degree of restriction obtaining over a range of possibilities. In the limit of maximum restriction, of minimum generality, reference to

[10] Cf. note 9, Bartlett, *ibid.*

the set of pure non-essential (contingent) features of a given phenomenon is possible. And, inversely, in the limit of minimum restriction, reference is possible to the set of pure essential (invariant) structures constitutive of the given phenomenon with respect to a class of phenomena of similar structure. The latter set expresses the " formal constitution " of the given phenomenon, while the former set refers to its " material constitution ".

Since an invariant structure is essentially related to a range of possible objects exhibiting that structure, the formal constitutive structures of a phenomenon are regarded as being intrinsically relative to a range of material possibilities. It is in this sense that phenomenology, as it is here proposed, is relative to a given, and is in this measure empirically bound. Reference to phenomena, which comprise the domain of possible experience, is fundamental to phenomenology.

C'. This approach to phenomenology may be termed 'transcendental' in that the fundamental concern is to elucidate the conditions which must be satisfied by objects of possible reference in order that reference to them may obtain. Such a foundational elucidation of referential systems attempts, then, to render explicit the structures upon which consistent forms of reference depend.

D'. I now turn to describe briefly a methodology developed to answer the need in phenomenology for an adequate theoretical framework in which transcendental problems of reference can be elucidated.

This methodology is intended to meet the need for a procedure which is capable of determining and correcting a form of invalid reference involved in conceptual misconstructions in foundation work in phenomenology, as well as in concepts basic to the sciences, concepts which phenomenology would seek to explicate in terms of an analysis of conditions of valid reference.

Phenomenological methodology, according to this view, is specifically interested in identifying, avoiding, and eliminating obstacles that stand in the way of the kind of scientific elucidation it seeks to develop. Phenomenology may therefore be characterized as committed to a model of " *explication as elimination* ", as suggested by Quine:

> We have, to begin with, an expression or form of expression that is somehow troublesome. . . . But it also serves certain purposes that are not to be abandoned. Then we find a way of accomplishing those same purposes through other channels, using other less troublesome forms of expression. The old perplexities are resolved [11].

[11] Willard V. O. Quine, *Word and Object* (Cambridge, Mass.: Technology Press 1960) p. 260.

The resolution of certain kinds of " perplexities " will be closely associated with the meaning of ' phenomenology ' here. In an obvious sense, a kind of " conversion of meaning " will go hand-in-hand with the resolution and elimination of difficulties that stand in the way of phenomenological explication.

The particular form of invalid reference that will be characterized here is termed a *' projective misconstruction '*, which can be understood in the following manner:

A phenomenon of any kind is relative to a determinable context of identification. In general, it is possible to characterize any phenomenon in relation to other phenomena or structures to which the phenomenon is essentially relative. In this sense, the propositions of a non-euclidean geometry are essentially relative to that particular system for their sense and truth-value. These patterns of relativity are to be found in every discipline; a network of relativistic relations constitutes or provides for the foundation for a discipline's internal unity.

Some interesting consequences follow from an analysis of these kinds of relations: it can be demonstrated that if two things are connected by a relation of essential relativity, then to affirm one out of connection to the other is logically inconsistent. As an example, consider a Cartesian coordinate system simply as a certain kind of system which provides definite means for identifying the position of objects in relation to an ideal origin in the framework. An object, the Cartesian coordinates for which are given, is represented in such a way that its position can be located in a Cartesian coordinate system. If these coordinates — without appropriate coordinate-transformation — are thought to locate the object from the standpoint of a Polar coordinate system, a conceptual misconstruction results. By a ' conceptual misconstruction ' I mean a ' logically invalid proposition resulting from an improper operation with a set of conceptual structures '.

The notion of a projective misconstruction, or, more simply, of a projection, is understood in the sense that reference to objects may be likened to coordination. Correct reasoning, or proper operation with a set of conceptual structures, presupposes valid coordination. In order to eliminate and to avoid this variety of improper reference or coordination, the *method of de-projection* is established. De-projection is a procedure with respect to which it is possible to clarify and restore the regulation of coordinations in accordance with the structural principles essential to the constitution of a given system. In the example, this would amount to showing that a position is appropriately designated in a Polar coordinate system if and only if the position is identified in a form complying with the understood conventions for specifying

points in that kind of coordinate system. This would of course require re-expression of the initial coordinates of the specified position through coordinate-transformation.

Relative to a given frame of reference, identifying references are essentially possible. An identifying reference is such that an ascription to that which can be the subject of an ascription establishes that what is ascribed and that that to which ascription is made, are one and the same. Such an ascription determines reference to that which is thereby identified such that the subject of the identification is fixed within a structure which allows for the possibility that the same subject can be re-identified.

An improper coordination results when reference obtains in a manner that does not conform to the syntactical organization of the frame of reference that conditions the possibility of the reference. A coordination then obtains which is improper in the sense that the reference itself does not satisfy what must be granted for it to be possible. When this invalid form of reference is explicitly described, it can be exhibited as devoid of sense and absurd.

Clearly, however, not all meaningless and inconsistent propositions express projections. A projection is a particular kind of coordination which must satisfy the following conditions:

(1) *A projection requires as a condition of its possibility that a phenomenon be disconnected from certain of its essential relations to the coordinating structure required for its possibility. In other words, there must be a severing of the essential relativity of the object of reference to its context of reference.*

(2) *The phenomenon must be asserted to be in certain respects autonomous of its context of reference. Reference must be made to the phenomenon in such a manner that denies or ignores one or more essential determinants of its contextual relativity. The coordination is projective in these respects.*

It must be understood that (1) and (2) are descriptions of conditions which must be satisfied by any projection from the standpoint of an analysis of its essential structure. It is not as if a projective misconstruction takes the form of an explicit severing of essential contextual relations — although this may be so in deliberately constructed cases. But this condition of explicit severing of essential coordinative relations must be satisfied once the projection is expressed in the form of an assertion that specific structures are *not* essentially connected to the context.

Let the term 'coordinate' be understood in the sense of a 'determinant of reference to that which can be the subject of an identifying reference'. A

coordinate is therefore identifiable within a coordinate system which assures it an appropriate context of reference. Any reference to a phenomenon establishes a relational system which conditions the possibility of that reference. In keeping with Quine's dictum " no entity without identity " [12], a phenomenon is intrinsically specifiable in terms of what renders it determinable — in the example, its " coordinates ".

Through a description of the constitutive structure of a phenomenon, de-projection retrieves to that phenomenon its coordinates which, as projected, are regarded as autonomous of the coordinate system(s) entailed by these coordinates. Thus, de-projection is concerned with elucidating the structures necessary for the possibility of the referential character of coordinates.

When misconstructions implied by a disregard for the constitutive relativity of a given system of reference are circumvented by heeding the conditioning principles upon which the possibility of the structures examined depend, analyses introduce, as it were, no supplementary content, or assumptions. As a method for accurate description, the formal structure of de-projection is tautologous — for, in making explicit the constitutive elements of that which is described, a point is reached where it is clear that the affirmation of a concept, or reference to a phenomenon, must at once involve the constitutive structures which guarantee the possibility of that concept or reference. It is precisely because de-projection is empty of content that it can authorize a transition from one formulation to another, while guaranteeing their equivalence, and without risking an automatic introduction of error.

The first condition above specifies that a phenomenon is to be considered as dislocated from its essential contextual relativity. The second condition specifies that this dislocation is to be formulated in the form of an assertion of the autonomy of the phenomenon with respect to its context of reference. The first condition denies the possibility of the phenomenon by separating the phenomenon from the context conditioning its possibility. After F. B. Fitch, such a denial is termed ' self-referentially inconsistent '. The separation involved is strictly speaking impossible, so long as reference is actually *understood* as intending a certain object of reference. However, *what* is said of that object of reference — namely, that it is autonomous of the context conditioning its possibility — comprises a self-contradictory and projective assertion. From the standpoint of de-projective analysis, the above separation must consequently be considered a *mistaken* separation.

The second condition, then, involves asserting the phenomenon while the grounds for its possibility are excluded. Thus, a projective misconstruction

[12] Leonard Linsky, *Referring* (New York: Humanities Press 1967), p. 27.

would attempt to disconnect two or more things which are essentially relative to one another, where this essential relativity of the one to the other is necessary for either to be possible [13].

The elimination of projections follows according to the coordinating principles of the context within which a projection obtains. Three distinct moments of analysis are made prior to the de-projective correction of a projective misconstruction. First, the constitutive coordinative structure of the concept or reference in question must be adequately described. This description must specify the essential restrictions imposed by this structure upon possible coordinations. Second, the assertion involved in the projective misconstruction must be explicitly formulated. The formulation will specify the nature of the " projective demand " by designating the nature of the asserted autonomy of the given concept or reference with respect to its conditioning context. Third, the opposition of the projective demand to the regulative principles of the context must be verified to result in a contradictory and meaningless formulation.

Together, these preliminary analyses render explicit the constitutive structure of the given context and demonstrate that reference to the concept or reference in question is at once necessarily relative to those constitutive elements guaranteeing the possibility of that concept or reference. De-projective analysis is completed through a reconciliation of the constitutive coordinate structure with the misconstruction which was originally in opposition to that constitution. This final phase of de-projection involves a correction of the projective coordination, imposing upon the coordination regulation according to the regulative structure of the context of reference, which in turn conditions the possibility of reference to the given concept or reference [14].

The idea of phenomenology as being closely tied to a model of explication as elimination reveals a similarity between the function of de-projection and the rôle of the *epoché* or reductions in Husserl's phenomenology. Where the latter isolates by bracketing or suspending a region of " dubitable nature ", projective misconstructions are eliminated in the former. In a rather limited sense, the phenomenological *epoché* may be thought to function at

[13] Up to this point, I have discussed the essential relativity of any object of reference to some context of possible reference. The inverse relation of essential relativity of any context of possible reference to some domain of one or more possible objects of reference is a consequence of the fact that "an invariant structure is essentially related to a range of possible objects exhibiting that structure". (Cf. above, p. 179.) For this reason, among others not treated here, a Platonism concerning possible systems is avoided.

[14] A full account of the notions of projective misconstruction and of de-projection is given elsewhere. See note 9.

times as a suspension of the absurd [15], and here Husserlian phenomenology shares with de-projective phenomenology if not the same, then at least a common, interest. An "essential residuum" is left in each case: in de-projective phenomenology, a de-projectively clarified field of phenomena; for Husserl, indubitable consciousness. At times these residua overlap; usually, however, they do not, and there the similarity to phenomenological reduction breaks down.

E′. De-projective phenomenology begins in the employment of a strict methodology, where certain explicit deviations from consistent forms of reference are considered invalid in the sense that the description formulated in a de-projective analysis must itself be granted as a correct description in order for valid forms of reference to obtain. De-projection leads, when this is possible, in its final corrective phase, to a re-formulation free from projective misconstructions of the initial concept or reference in question.

The description of the constitutive structure of a phenomenon or group of phenomena from the standpoint of a given framework cannot be universally generalized. The description is relative to a given framework, and can be considered invariant only in relation to an isomorphic system of frameworks. For example, the translatability of a proposition is always relative to systems providing adequate means for the expression of that proposition. It is merely factual that not all systems have equally adequate means of expression.

Phenomenology undertaken in this manner is primarily interested in investigating the essential structure of individual phenomena taken in relation to definite and general classes of phenomena in which they have membership. From this standpoint, a de-projective phenomenological description of a given phenomenon tends to minimize the non-essential, *individuating* features of that phenomenon. As such, de-projective phenomenology is in the nature of a *generalized analysis,* as developed by A. A. Gukham [16], the task of which is to elucidate the formal structure of a group of phenomena with respect to which the essential structure of a given phenomenon is covariant. De-projective phenomenology, incorporating a transcendental theory of reference, can be characterized as a descriptive science on the level of maximum theoretic generality, the aim of which is to render *explicit* the structure of possible objects of reference, the structure of possible experience.

[15] See above, p. 5 (i).
Log. Unt. II. 2 v § 16, § 27; III Appendix § 5; *Ideen* I § 48; S. Bachelard, *La Logique de Husserl* (Paris: Presses Universitaires 1957), p. 136.

[16] Aleksandr Adolfovich Gukham, *Introduction to the Theory of Similarity,* trans. ed. Robert D. Cess (New York: Academic Press 1965).

III

The purpose of reflective philosophy is to elucidate, explicate, or disclose the *implicit* structure of possible, or merely actual, experience. For here, it is said that " to analyze is to explicate the implicit " [17]. The development of reflective philosophy may be understood as a growing consciousness of the nature of the primary task of philosophy: to render the implicit explicit. If phenomenology is committed to " radical self-understanding ", as Husserl would put it, then it is important that phenomenology account for itself in the terms of the tradition to which it belongs. Therefore, it will be useful (1) to consider what place should be accorded to the terms ' explicit ' and ' implicit ' in the formulation of a reflective philosophical approach, and then (2) to take note of the relation between the meaning established for these terms and that of ' reflective experience ' and ' pre-reflective experience '.

Before proceeding in this manner, it will be helpful and of interest to refer here to several passages in Husserl's *Ideas* I, which will provide a background for the discussion here.

[§ 35] " Every perception of a thing has a zone of background intuitions (or background awarenesses . . .) " such that every perception of some thing shades off into a total context, a " co-perceived objective ' background '. "

" . . . it is here implied that certain modifications of the original experience are possible, which we refer to as a free turning of the . . . *' mental look '* . . . from [e. g.] the paper at first *descried* to objects which had already appeared before, of which we had been ' implicitly ' aware, and whereof *subsequent* to the directing of one's look thither we are explicitly aware, perceiving them ' attentively '. . . . "

" . . . we know that it is the essence of all such experiences . . . to exhibit that remarkable modification which transfers consciousness in the *mode of actual orientation* to consciousness in *the mode of non-actuality* and conversely. At the one time the experience is, so to speak, ' explicitly ' aware of its objective content, at the other implicitly and merely *potentially.* "

[§ 78] " . . . *every variety of ' reflection '* has the character of a *modification of consciousness.* . . " The " unreflective experience-datum undergoes a transformation — into the mode, that is, of reflective consciousness (consciousness of which we are aware). "

[17] Paul Ricoeur, *Husserl: An Analysis of his Phenomenology,* trans. E. G. Ballard and L. E. Embree (Evanston: Northwestern University Press 1967), p. 99

[§ 79] " . . . phenomenology . . . makes . . ., as a fundamental condition of its possibility, positive affirmations concerning unreflective experiences. These it owes to reflection, or, more accurately, to reflective intuition of the essence [18]. "

(1) When " implicit reference " is made to a phenomenon, the following formal conditions (or some other conditions which are reducible to these) must be satisfied:

From the standpoint of a system K, it is possible identifiyingly to refer to sub-systems I and J, where systems I and J have a given common structure. System-J, though sharing a common structure with system-I, is such that a set of elements, A, B, C, are associated with the system as constitutive of it. Since the sub-systems have a common structure, from the standpoint of system-K, A, B, C can be correlated with system-I, even though these elements are directly associated only with system-J.

Insofar as the correlation holds, A, B, C are said to be both *explicit* constitutive elements of system-J, and *implicit* constitutive elements of system-I. To generalize, the implicit has a structure conditioning the range of possible explicit structures which can validly be correlated with it.

System-K, then, provides a possible framework within which the relation between the terms 'explicit' and 'implicit' is determinate. A projective misconstruction obtains if an object of reference is characterized both as (a) having an " implicit " structure of a certain sort, and as (b) having such a structure out of connection to such a system K. To say of an object of reference that it has a certain character implicitly, while reference to systems like J and K, to which predication of the term 'implicit' is necessarily relative, is denied or neglected, — to maintain this, immediately gives way to an inconsistent and absurd misconstruction. Such a misconstruction will be termed a 'projection of the implicit'.

Reflective philosophy may accordingly be viewed as posing the general task of explicating any subject-matter by validly correlating an explicit description with that subject-matter. Some framework which permits reference to such a correlation is necessarily presupposed whenever reference is made to an " implicit content ". A projection of the implicit obtains if such reference is made in apparent autonomy of the presupposed framework. Any reference to what is " implicit ", independent of a framework permitting a correlation between what is implicit and what is expressed in an explicit description, manifests a projective misconstruction of this kind.

[18] Cf. also *Ideen* I, §§ 36, 69, 77.

(2) It is clear with some reflection that only from the standpoint of an embedment-system of higher order, can reference be made to an egologically modified or affectively modified phenomenon (e. g., a " desired apple ") so as to permit the discrimination of a specific egological modification (the " desire ") with respect to the phenomenon (the " apple ") which it modifies, or with which the modification is correlated. It should be emphasized that the distinguishability of a specific egological modification, or attentional character, in relation to the phenomenon so modified, is essentially relative to a context of reference which provides for recourse to a higher order embedment-system [19]. In relation to a context which does not provide for reference from the standpoint of such an embedment-system, a phenomenon and the attentional character modifying it cannot be distinguished.

A context the structure of which does not provide for recourse to a higher order embedment-system is said to comprise a " pre-reflexive standpoint ", in contrast to a " reflexive standpoint ", in relation to which such an embedment-system is established [20]. The former is said to be " pre-reflexive " since it frequently is possible to evidence a correlation between two given phenomena, where one phenomenon temporally precedes the other and may or may not be egologically modified, while the second is explicitly modified, that is, is such that any egological modification can be distinguished from the phenomenon modified. This correlation may not be generalized, however, inasmuch as either phenomenon in question may be given in contexts without the other. The distinction here between the two phenomena is a simple expression of the difference, as it were, between phenomena and their explicit description.

From a reflexive standpoint, then, it is possible for the purposes of descriptive analysis to differentiate between a given phenomenon and an attentional character which may modify that phenomenon. It is emphasized, once again, that such a distinction is essentially relative to that reflexive standpoint. It is therefore projective to " carry over " the results of a reflexive analysis of phenomena to phenomena which are thought to be " pre-reflexively constituted " independently of the very framework in terms of which the notion of " pre-reflexive constitution " is applicable. Thus, reference to a

[19] *Log. Unt.* II. 2 v § 23.
It follows that it is projective to represent an attentional character autonomously of such a context.

[20] The terms 'pre-reflexive' and 'pre-reflective', 'reflexive' and 'reflective' are distinguished here in that the first term in each pair is associated with general structures of systems of possible reference, whereas the second term in each pair is a familiar occurrence in the literature pertaining to "consciousness".

" pre-reflexive context " from a reflexive standpoint can only be understood, *simultatione,* to involve a " pre-reflexive context " [21].

The projections distinguished under (1) and (2) above are variously represented in the passages quoted from *Ideas* I; it is possible here to enumerate only two of the assertions assumed there which can be shown to involve projections.

(a) Reference from the standpoint of a reflective framework involves a " modification " of a pre-reflective content, which remained unchanged until it was identifiyingly referred to from that standpoint.

(b) Descriptions are possible of the constitution of what is implicitly given without reference to a reflective standpoint.

The corrective phase of de-projection would yield the following re-formulation: Phenomenology as an explication of the implicit, or as a foundational elucidation of possible objects of reference, is so constituted itself, as an approach to a field of problems, that any phenomenological description is intrinsically relative to the phenomenological framework. De-projective phenomenology, which is elaborated in terms of a transcendental theory of reference, may obtain certain descriptive results concerning such notions as " pre-reflective experience " or " the implicit ", but these results cannot be taken out of relation to frameworks rendering those results possible.

It is difficult to understand why phenomenologists have been unaware of this " limitation " placed upon their activities, while the same " limitation " has been accepted and understood by mathematicians for centuries. A geometer would not claim any validity for his results outside of the system(s) in terms of which the validity of these results can be posed as a question.

Prof. S. Bartlett
Research Fellow
Max-Planck-Institut
zur Erforschung der Lebensbedingungen
der wissenschaftlich-technischen Welt
Starnberg, West Germany

[21] It may be hazarded that perhaps only "artistic involvement" permits wholly intrinsic representation of the essential structure of pre-reflexive phenomena.

CHAPTER 11

Self-Reference, Phenomenology, and Philosophy of Science

SELF–REFERENCE, PHENOMENOLOGY, AND PHILOSOPHY OF SCIENCE

by

STEVEN J. BARTLETT
Saint Louis University
(U.S.A.)

Abstract

The paper begins by acknowledging that weakened systematic precision in phenomenology has made its application in philosophy of science obscure and ineffective. The defining aspirations of early transcendental phenomenology are, however, believed to be important ones. A path is therefore explored that attempts to show how certain recent developments in the logic of self-reference fulfill in a clear and more rigorous fashion in the context of philosophy of science certain of the early hopes of phenomenologists. The resulting dual approach is applied to several problems in the philosophy of science: on the one hand, to proposed rejections of scientific objectivity, to the doctrine of radical meaning variance, and to the Quine-Duhem thesis, and on the other, to an analysis of hidden variable theory in quantum mechanics.

Phenomenological philosophy began in rigor and has gradually submitted to imprecision. Early in its development, phenomenology was cultivated in close connection with natural science and mathematics, and was inspired by an appreciation of exact standards of justification.[1] On the whole, it seems evident that phenomenology

1. J. Robinson (1788), W. Whewell (1847), E. Mach (1894) variously conceived of phenomenology as a methodological tool of research in physics. Brentano (1888) extended

has placed this attitude to one side, and has become a humanistic tool of interpretation currently under the wing of hermeneutics and existentialist thought. Even within the individual lives of its main contributors, there has been a perceptible transition from scientific standards of exactness to humanistic *Verstehen*.

As a result of this change of orientation, phenomenology offers what is often judged to be an obscure and terminologically top-heavy set of tools for use by philosophers of science. However, phenomenology, at least in the earlier thought of Husserl and to a lesser extent, in some works of Meinong and Brentano, offers a methodology which is distinguished by a number of properties of special interest to philosophy of science.

I

Phenomenology as conceived by the young Husserl – and I have in mind that variety of phenomenology which identifies itself as non-genetic (non-explanatory), descriptive, transcendental phenomenology – aspired to these ends: It sought to provide a method of descriptive analysis capable of explicating the transcendental preconditions which of necessity would need to be satisfied in order for it to be possible for certain objects of conscious life to possess essential properties which they do. An easily identifiable Kantian thread bound together a variety of interests in studies of the constitution of particular objects of consciousness, the constitution of the ontology of regions, the constitution of time, etc. In these investigations, phenomenology was to comprise, in the words of Stumpf, a "neutral pre-science" (*Vorwissenschaft*) which would introduce into its framework of descriptive analysis no special presuppositions,

Whewell's classificatory conception into phenomenological psychology. Baron Jakob Johann van Uexküll (1909) published a group of studies undertaken from a phenomenologically sensitive ecological standpoint, well ahead of his time. Husserl's doctoral research under Weierstrass on the calculus of variations supported his Habilitation thesis on the concept of number (reworked later into the uncompleted *Philosophie der Arithmetik* of 1891). Husserl's *Logische Untersuchungen* (the first volume published in 1900) and his *Formale und transzendentale Logik* (1929) add to this early picture of phenomonology's close association with the sciences.

and would enable the phenomena treated to speak for themselves, as it were, without suffering from perturbations due to the method employed in their analysis. As a *radical* enterprise, in the special phenomenological sense of this term, this presuppositionfree approach would seek to account for its own transcendental structure. It would, that is to say, possess the property of self-reflexiveness, falling within the scope of its own proper subject-matter.[2]

The methodology resulting from this rigorous phenomenological orientation can be distinguished, then, by its claims to presuppositionlessness and self-reflexivity, and by its transcendental concern to explicate preconditions which must be granted for individual phenomena, classes of phenomena, and a wide range of properties and relations between them, to be possible.

Such a proposal, had it borne fruit, would have found important applications in the context of a study of scientific theories. Ideally it would have provided a wholly *intrinsic* mode of analysis of the structure of a scientific theory, because it would have comprised an approach that claimed to impose no external standards of criticism. The results of such an intrinsic critique of a scientific theory could indeed "command the assent of all who are competent to form an opinion."[3] Such a phenomenological approach would make possible an analysis of the presuppositional structure of a theory if not in its own terms, then *of* those terms from a neutral standpointless metraframework compatible with the framework of the theory. The approach would constitute a rigorous metatheory which could be applied in the dispassionate spirit of scientific neutrality both to individual scientific theories as well as to itself.

This proposal – and there is no judgment made here of the in principle possible future the proposal could have had or may yet enjoy – historically has not been successful in the context of scientific interest. This is not, indeed, the sole arbiter of a philosophical methodology, but it is the one of interest to philosophers of science who share the desire to free their discipline from the uncertainties

2. For a fuller account of this interpretation of early phenomonology, see Bartlett (1975).
3. Russell (1914), p. 69 in 1972 edition.

of controversy, and to contribute to the development of what has been called a "vertical discipline", one which builds progressively upon the demonstrated results of the past.

II

Frederic Brenton Fitch, a mathematical logician with an unusual sensitivity to things philosophical, has proposed an approach to philosophy somewhat analogous to the transcendental phenomenological variety I have, perhaps too summarily for some, laid to rest. The "universal metalanguage for philosophy" that Fitch has endeavored to describe bears a close resemblance to one of the defining properties we have mentioned in connection with the methodology of rigorous, scientific phenomenology.[4] Fitch's universal metalanguage has not been formulated so as to include the critical resources needed to make possible its application as a tool of criticism by philosophers of science. Yet, unlike the approach of transcendental phenomenology, the view is clear, and with some phenomenologically-motivated supplementation which I shall suggest, appears to lend itself extremely well to certain of the objectives of philosophy of science.

Fitch argues that the level of generality required for much philosophical discourse is such that the Russell-Whitehead theory of types must be rejected. Philosophical discourse desires "extreme comprehensiveness" of the kind which requires self-reference. In philosophy, this situation is frequently encountered:

> Theories are constructed which purport to deal with all entities whatsoever and which therefore have an unrestrictedly extensive subject matter. In dealing with all entities, such theories in particular deal with all theories, since theories are themselves entities of a special sort. In philosophy we thus encounter theories about the general nature of theories ...
>
> If a theory is included within its own subject matter, we say that it is a *self-referential* theory.[5]

4. To be precise, Fitch discusses a *family* of languages any one of which avoids Tarski's limitative criterion for an acceptable definition of truth.
5. Fitch (1952), p. 218.

In particular, the concern of phenomenology "to explicate its own foundation" requires the self-reflexiveness which characterizes a self-referential theory.

Discoveries since the turn of the century of set theoretical, semantical, and pragmatical paradoxes rendered suspect any self-referential theory of this sort. Self-reference was blamed, and it was banned by the cures that were prescribed to eliminate the occurrence of paradox. In the process, and virtually ignored by the phenomenological community, the Cartesian radicalness of phenomenology was made incapable of realization. The road to the desired self-reflexiveness of the phenomenological approach would remain closed as long as it could be proved that such a theory of theories, or science of sciences, was paradox-generating.[6]

The disturbance due to the discovery of the paradoxes was felt by another field of study, within philosophy of science. Philosophy of behavioral science has often sought the extreme degree of comprehensiveness Fitch describes. A philosophical reflection on human behavior comprises, when undertaken by a human being, a human behavior which falls within the scope of concern of behavioral science and its philosophy.

Similarly, a comprehensive theory of human reflection, when the theory itself is an expression of this capability, requires self-reference.

The anti-paradox cures which were prescribed and which have almost universally been endorsed (e.g., variations on the theory of types and Tarski's limitative semantical results), effectively blocked hopes for extreme comprehensiveness involving self-reference.

Fortunately, in the years since paradox paranoia first disabled the logic of self-reference, certain constructive attempts were made to save the self-referential interests of phenomenology, philosophy of behavioral science, studies of human reflection, etc. In 1963, Fitch demonstrated that non-Tarskian systems do exist which (a)

6. It would be possible to escape this conclusion if it could be shown that the methodological framework of phenomology forms a system of an essentially non-formalizable kind, to which formal set theoretical, semantical, and pragmatical constraints do not apply. However, this has not, as far as I know, been done.

are provably consistent, and (b) permit self-reference.[7] Others, including Smullyan, Myhill, and R.M. Martin, have reinforced Fitch's general conclusion.

As a result of these and similar efforts, it is no longer necessary to avoid all forms of self-reference in order to avoid the occurrence of paradox nor is it necessary to resort to an endless ladder of formal metalanguages. The extreme comprehensiveness desired by much philosophy, by phenomenology, and by other fields, may now again be viewed in a favorable light.

III

With these traditional formal blocks removed, it is possible to consider how a self-referential universal metatheory may be constructed as phenomenology wished. Certain of the fundamental interests of rigorous phenomenology can perhaps be realized in a more perspicuous and more effectively applicable form, following recent contributions to the logic of self-reference.

Specifically, (1) phenomenology's wish to explicate the essential structure of phenomena in a manner free from special presuppositions may be paired with (1′) the intrinsic style of self-referential criticism of which a number of accounts are now available. (2) The self-reflexivity of transcendental phenomenology has a real analog in (2′) a self-referential metalogic that seeks to identify preconditions of referring. (3) The twin foci of phenomenology's intentional and transcendental forms of analysis may be paired with (3′) these two similar foci: a pragmatical description and analysis of intentions involved in referring, and a metalogical account of referential presuppositions subscribed to. Finally, (4) the wish in phenomenology for non-controversial results may be fulfilled by (4′) the proof-oriented approach of a self-referential metalogic of reference.

In earlier work, I have explored the idea of a general metalogic of reference, and have examined certain of the formal properties of

7. Specifically, these systems permit semantical self-reference, which is needed for such a system to formalize its own truth concept. Cf. Fitch (1963).

the resulting metalogic.[8] Here I would like to consider the concerns of a general metalogic of reference which correspond to analogous phenomenological interests.

A sense of presuppositionlessness is achieved by intrinsic, self-referential criticism of a position. Henry W. Johnstone, Jr., has attempted to show that philosophical arguments are successful only when ultimately they are *ad hominem*. For Johnstone, valid critical argument in the *ad hominem* style takes seriously claims made within the framework of a position, and then shows how some claims are self-refuting, short-circuiting the intended purposes of the advocate of the position.

Johnstone distinguishes seven types of philosophical argument. One of these, which he calls 'the charge of denying presuppositions', is worth mentioning here.[9] A denial of presuppositions occurs when a statement made on behalf of a position denies just what the position presupposes. As an example, Johnstone gives the statement, "life is a dream," which is meaningful only if it is presupposed that a meaningful distinction between dreams and waking-states is possible. But this possibility is precisely what is denied by the statement.

Since philosophical argument appears to serve primarily a critical function for Johnstone, *negative* disputation is emphasized by him. (So it was when Kant suggested, in a 1772 letter to Lambert, the need for a *phaenomenologia generalis*, a "negative science" propaedeutic to metaphysics). An approach resembling the one suggested by Johnstone can, however, be used equally to show, as we shall see, the reverse: that one cannot *not* accept certain claims made within the framework of a position.

John Passmore has formulated a position similar in some respects to Johnstone's. Passmore reviews three ways in which one can "con-

8. The latter study will appear in a forthcoming paper, "Referential Consistency as a Criterion of Meaning"; the former may be found in Bartlett (1970); (1975); (1976); and (1978) §§ 10, 12.

9. Cf. Johnstone (1959), pp. 90f. It may later be noticed that Johnstone's denial of presuppositions, if extended beyond its intended factual, *ad hominem* range of application, closely resembles the metalogical variety of self-referential inconsistency. (See below.)

tradict oneself'. One of these results in what Passmore calls an "absolute self-refutation". It resembles Johnstone's denial of presuppositions. In Passmore's case, however, it is not that the special presuppositions of a particular position are denied, but "implicit assumptions ... about the conditions of inquiry." These "invariant conditions of discourse" cannot coherently be repudiated. Attempts to deny these conditions result in absolute self-refutation. For example,

> ... it is presupposed in all discourse that some propositions are true, that there is a difference between being the case and not being the case, and to deny this in discourse is already to presume the existence of the difference – since otherwise the notion of 'denying' is quite meaningless ...
>
> Only if a philosophical argument can show in this way that a sentence can propose nothing – because what it asserts, if it were taken to propose something, would be inconsistent with the presuppositions of all proposing – is it pointing, I suggest, to an absolute self-refutation.[10]

Although the positions articulated by Johnstone and Passmore complement one another, there is disagreement. Johnstone, for example, does not accept the view that Passmore's allegedly absolute self-refutations cannot be evaded. Johnstone agrees that "invariant conditions of discourse" do exist and are significant in the context of self-refutations. But, he argues,

> I only insist that we think of such invariant conditions as being hypothetical rather than categorical in form. While I am suspicious of 'Every sentence conveys something', and doubt it has a role in self-refutation, I would be perfectly happy with 'If a sentece is used as an assertion, it must convey something'. For I am willing to see the consequent of this conditional apply to every sentence to which the antecedent applies. It is only the cases to which the antecedent does not apply that cause me to reject the categorical version.[11]

For Johnstone, an effective argument must always take into account the intentions of the advocate of the position under analysis. For Passmore, this is not always necessary because some presuppositions of discourse cannot be suspended by personal fiat. There are other disagreements in the extensive literature treating self-refutation and *ad hominem* argumentation, but they need not concern us here.

Neither Johnstone nor Passmore has shown that the invariant con-

10. Passmore (1961), p. 68.
11. Johnstone (1964), p. 478.

ditions of discourse which both authors claim exist, do exist. A few examples are given, but for most purposes these illustrations fail to establish the general thesis.

The mathematical logician Paul Lorenzen has also endorsed undeniable conditions of discourse in his treatment of "*elementary sentences*", which can be used to express basic assertions and denials. He reasons that the

> ... decision to accept elementary ways of speaking is not a matter of argument. It does not make sense to ask for an 'explanation', or to ask for a 'reason'. For to 'ask' for such things demands a much more complicated use of language than the use of elementary sentences itself. If you ask such questions, in other words, you have already accepted at least the use of elementary sentences.[12]

Collingwood and his constructive interpreter, Rynin, also argue that there exist "absolute presuppositions" which, although not themselves truthfunctional propositions, underlie as necessary conditions for systematic thought propositions that *are* true or false. For Collingwood, a study of absolute presuppositions is a central task of philosophy. Such a view of philosophy requires self-reference.

> Philosophy is reflective. The philosophizing mind never simply thinks about an object, it always, while thinking about any object, thinks also about its own thought about the object.[13]

IV

We have described several views concerning self-refutation which are of interest in the context of an approach to intrinsic, and, in some as yet undeveloped sense, presuppositionless analysis. The views we have reviewed share a self-referential perspective, and focus either on (a) what must be presupposed as a general condition of discourse or of systematic thought, or on (b) what the advocate of a position in fact is forced to acknowledge if his intentionsare to be

12. Lorenzen (1969a), p. 14.
13. Collingwood (1946), p. 1.
Overtones of self-reference are found, too, in Lorenzen's claim, in connection with his operative logic, that "the method is identified with its own result." (1969a), p. 89. Cf. Lorenzen (1969b).

realized. Whichever alternative is followed, the claim is made that the conclusion of an argument by means of self-refutation is not dependent upon the prior acceptance of special norms or criteria alien or external to the position analyzed. Analysis of this kind uses, so to speak, the energy of a position to provide a critique of that position. Philosophical argument in this style suggests a form of intellectual judo. In this general sense, it advances no special presuppositions of its own, endorses no partisan criterion of meaning, but has what we might be tempted to call a "tautological structure": A formulation of the regulative metalogic followed is devoid of positive content, and would articulate general principles that express equivalences of meaning.[14]

In a second analogy to phenomenology, transcendental self-reflexivity corresponds in a self-referential metalogic to a concern to identify preconditions of referring. A metalogical precondition of referring is specified when any attempt to reject that condition results in self-referential inconsistency. This "test" lends itself to formalization and supplies an intrinsic analysis with a logically non-arbitrary and compelling criterion,[15] as I shall try to illustrate. Furthermore, such a critical criterion complements Johnstone's approach to a denial of presuppositions, and is in agreement with Fitch's understanding of a presupposition as "an assumption whose denial is self-referentially inconsistent."[16] A metalogical precondition of referring is "absolute" within all contexts of reference of a certain kind. It will, as things turn out, share some of the properties ascribed by Passmore to his invariant conditions of discourse, and some of those ascribed by Johnstone to his *ad hominem* approach to philosophical argument.

Two major varieties of self-referential inconsistency have been

14. Barlett (1970), Chapter 1.4.

15. This is shown in the forthcoming paper mentioned in note 8.

16. Fitch (1952), p. 221. Fitch has in mind here that the acceptance or rejection of accepted principles of logic must rely upon the use of at least some of these principles. The kind of self-referential inconsistency he has in view turns out to be of a lower "modal order" than the transcendental variety to be described: that is to say, Fitch is concerned with principles which *in fact* must be presupposed, in contrast to presuppositions which *in principle* cannot be rejected.

studied in analyses of self-refutation. It will be important to us to distinguish these clearly, since, in the transformation of exact phenomenology to a metalogic of reference which I am suggesting, these two varieties comprise rough analogs of the phenomenological modes of analysis, intentional and transcendental. In the remainder of this section, we shall look at one of these, and discuss two divergent conceptions of presupposing with which it has been associated. The self-referential analog to transcendental analysis will be considered in the next section.

Both Passmore and Johnstone appear, in spite of their disagreements, to have in view fundamentally the same variety of self-referential inconsistency. Passmore claims that a proposition is absolutely self-refuting if the assertion of that proposition is equivalent to asserting both that proposition and its negation.[17] He gives a quite different formulation a few pages later when he claims that a proposition is absolutely self-refuting if it is taken as proposing something and if what the proposition does propose is "inconsistent with the presuppositions of all proposing."[18] The first claim has the form

$$p \text{ is self-refuting if } \vdash p \equiv p \,\&\, -p, \tag{1}$$

while the second has the form

$$\begin{aligned} &p \text{ is self-refuting if } (p \text{ proposes } q)\ \&\\ &(q \text{ is inconsistent with every } \alpha \text{ where } \alpha\\ &\text{is presupposed by all propositions}). \end{aligned} \tag{2}$$

It is not at all clear that (1) and (2) say the same thing, nor is it clear, given the confusion consequent to the array of analyses that have been supplied which treat the relation of presupposing, whether or not α should be interpreted as truth-functional.[19]

Johnstone's corresponding view is this: He argues that a valid philosophical criticism (a) identifies an inconsistency between an opponent's thesis and what the thesis presupposes, and (b) shows why one's opponent must acknowledge this inconsistency.[20] It

17. Passmore (1961), p. 60.
18. Passmore (1961), p. 68.
19. On this question, see, e.g., the controversy between Donagan (1962) and Rynin (1964).
20. Johnstone (1961), p. 353.

hould be clear that Johnstone's attention is focused on the *inten-ions* of his opponent. It is relative to an opponent's acknowledged ntentions that both (a) and (b) above are to be accomplished.

Both Passmore and Johnstone, while clearly not in total agreement, are concerned with what is *in fact* presupposed by the claims of a position. Passmore wishes to make recourse to invariant and ategorical conditions of discourse; Johnstone is more modest, contenting himself with "a *logic of intentions*"[21] revealed in a case-by-case analysis through the means of explicit controversy.

The variety of self-referential inconsistency of concern to both Passmore and Johnstone has been termed 'pragmatical' or 'performative'. A substantial literature has been devoted to its study.

A pragmatical self-referential inconsistency may be generally defined as follows:

> If a proposition p is used in a manner such that reference is made by an individual a to an object o at a place-time s, and if o is a pragmatical (or performatory) aspect of the use made of p by a at s, then p is called *pragmatically* (or performatively) *self-referential*. If a pragmatically self-referential proposition p is such that o falsifies p, then p is said to be *self-refuting*. (3)

The assertion, for example, "There are no truths", is self-refuting. It is absolutely self-refuting for Passmore in that "to assert is to assert to be true."[22] It is self-refuting for Johnstone if we can determine that the claim *is* intended by its propounder as a claim to truth (and is not, e.g., for him merely a sequence of meaningless noises or marks). In either case, the self-refutation concerns a factual aspect of the use made of a proposition. We note, then, that pragmatically self-referentially inconsistent or self-refuting statements are factually self-falsifying.

Such a pragmatical variety of self-referential inconsistency, if it is to be used in a non-paradox-generating context, appears to require the rejection of excluded middle.[23] The effect of this is twofold: First, a strengthened case for Strawson's familiar definition of 'presupposing' can be made. Strawson's view, that S presupposes S′ iffS is neither trueor false unless S′ is true, was objected to by

21. Johnstone (1959), p. 120.
22. Passmore (1961), p. 68 (Passmore's emphasis).
23. Fitch (1963) and (1952).

Rynin,[24] as follows: Rynin reasoned that if S presupposes S′, then both $S \rightarrow S'$ and $-S \rightarrow S'$ will be the case. By excluded middle, the conclusion follows that S′ is true, i.e., that all presupposed statements are true – which is of course highly doubtful. Rynin's objection is dissolved when excluded middle no longer applies. Strawson's analysis is left if not in a wholly unproblematical condition, at least repaired.

Alternatively, the rejection of excluded middle makes it intelligible to consider presuppositions in Collingwood's sense: For him, absolute presuppositions are neither true nor false, but they express what might be called "quasi-propositions" that articulate basic conceptual commitments.[25] This is the path I will pursue for reasons that will be evident shortly. It will be useful to make this restriction:

For purposes here, S is said to presuppose S′ in a frame of reference F iff S is neither true nor false unless S′ expresses a framework constraint that holds or is in force when S is asserted relative to F. According to this formulation, it makes no sense to say of a presupposition that it is true (or false) relative to a frame of reference, just as it makes no sense to say in the context of a game (e.g., chess) that a rule (e.g., the rule governing castling) is true (or false). It *does* make sense to speak of such rules as holding or as having been broken in a particular game, just as it is intelligible to say that a presupposition holds or is violated in relation to a claim made in a particular frame of reference.

When a presupposition holds or is in force, one may conclude that the consequent of an associated conditional is true. For example, a presupposition of referring to an individual named 'Rima' is that there exist some object of reference so named. This presupposition of name-use, when in force, implies that the statement "There is some object of reference named 'Rima' " is true. But it is a mistake,

24. See note 19.

25. Collingwood limited the term 'proposition' to what may be understood as the (true or false) answer to a particular question. He did not wish to view absolute presuppositions as expressing genuine propositions, since they are not answers to particular questions, but rather underlie the asking of such questions.

from the point of view described here, to equate the presupposition in question with the truth of the latter statement. The distinction made here takes into account differences between a rule, instances which satisfy it, and statements about those instances.

V

Referential presuppositions analyzed in this way, constitute, in the phenomenological sense, preconditions of valid reference. Their rejection, relative to a particular frame of reference, leads to a form of self-referential inconsistency which elsewhere I have termed '*projective*'.[26] A logic which studies relations between the referring use of concepts or expressions, and the referential preconditions which must be satisfied for that use to be meaningful, I have called a '*metalogic of reference*'. Its focus is, in the proper sense of the word, transcendental, and its range of concerns parallels that of transcendental phenomenology.

The strength of such an approach lies in the fact that the principles of the metalogic "*self-validate*" in the sense that their rejection leads to *projective* self-referential inconsistency. This metalogical variety of self-referential inconsistency is essentially distinct from the pragmatical variety. Where Passmore and Johnstone are alternatively concerned with absolute self-refutation or *ad hominem* argument in the context of factual conditions of discourse and acknowledged intentions, a metalogic of reference investigates the transcendental logic underlying all referring. Its interest is in preconditions of possible reference, and hence comprises a study which is properly metalogical.

The metalogical variety of self-referential inconsistency may be defined as follows:

> A proposition p is termed *metalogically self-referential* if p is such that (i) if p is asserted, reference is made by some individual a to an object o at a place-time s, and (ii) such reference metalogically presupposes endorsement by a at place-time s of a precondition M_p which must hold in order for p in principle to have a significant truth-value.
>
> If p is metalogically self-referential and p is such that its assertion denies one or more conditions which must be satisfied in order for it to be possible meaningfully to assert p, then p is said to be *projective*. (4)

26. Bartlett (1970), (1975), (1978).

A "precondition of reference", M_p, may be viewed as expressing a quasi-proposition, as described earlier. Such an M_p comprises a necessary condition of possible reference, a constraint which if violated in a particular context of reference results in projection.

Elsewhere I have argued that metalogical referential consistency constitutes a transcendental criterion of meaning in the sense that rejection of projective self-referential inconsistencies is a necessary condition of the possibility of meaning, truth-functionally understood.[27] From this point of view, a pragmatical analysis describes what one is *in fact* committed to in making an assertion, while a metalogical analysis describes what one *must* be committed to if an assertion *in principle* is to be meaningful.

A comparison of definition (4) and the earlier definition (3) of the pragmatical self-referential variety enables the reader to note these differences between the two forms of self-referential inconsistency we have discussed. The distinction between the two roughly parallels, I have suggested, the distinction between certain intentional and transcendental phenomenological analyses. On the one hand, a metalogical explication of preconditions of referring has an unmistakable transcendental orientation. On the other, *ad hominem* argumentation, or argumentation which attends to invariant conditions of discourse, requires a careful phenomenological description of identifiable intentional relations, either acknowledged by an individual advocate of a position or of necessity subscribed to in any use of discourse. A descriptive, intentional analysis of this kind would correspond closely to Johnstone's "logic of intentions" and to Passmore's study of absolute commitments of discourse.

We turn now to several examples which illustrate applications of this self-referential, phenomenologically-motivated metatheory to certain problems in philosophy of science.

VI

Carl R. Kordig has argued forcefully that most contemporary philosophies of science are self-referentially inconsistent in the sense of

27. See note 15.

being self-falsifying. His analyses emphasize the pragmatical mode of criticism, and merit attention.

For example, Kordig argues that the denial of objectivity in science and the doctrine of radical meaning variance are both self-referentially inconsistent. Specifically, both constitute self-falsifying assertions. The falsity of each claim is derivable from the assumption of its truth.

In connection with Kuhn's and Feyerabend's rejections of scientific objectivity, Kordig is in agreement with Scheffler: "Objectivity is presupposed by any statement which purports to make a cognitive claim. To put forth any such claim in earnest involves a presuppositional commitment to the view that the claim has an objective truth value."[28]

Kording opposes the views of Feyerabend (1962), Hanson (1958), Hesse (1963) and (1968), Kuhn (1962), Smart (1953), and Toulmin (1961) who have each argued that a shift from one scientific theory to another involves an incommensurable change in the meanings of the terms used, and hence that there can be no statements whose meaning is invariant across scientific theories. Kordig supplies an argument resembling Scheffler's: A statement which rejects radical meaning invariance is intended by its advocates to express the sort of meaning invariance it denies. Thus, its falsity follows from its assumed truth.

A possible objection is foreseen by Kordig: that the proposed rejection of objectivity in science and the endorsement of radical meaning variance are made from a restricted standpoint which is excepted from the claims made. It is true that in so doing the pragmatical self-referential inconsistency is evaded. However, the consequences of the evasion are unfortunate. The denial of scientific objectivity and the doctrine of radical meaning variance then result, according to Kordig, in an unjustified dualism: On the one hand, both scientific objectivity and invariance across scientific theories are denied; on the other hand, objectivity and meaning invariance are presumed in the special perspective of philosophy of

28. Scheffler (1965), p. 21.

science. This preference and privilege are not justified. Therefore Kordig is able to conclude that objectivity and meaning invariance in science cannot consistently be rejected, or this rejection entails the arbitrariness of dogmatism.

A third illustration of the pragmatical variety of self-referential argument is available in an analysis, also due to Kordig, relating to the so-called Quine-Duhem thesis. Quine (1963) and (1972) has been responsible for extending Duhem's thesis concerning physical hypotheses to all hypotheses. Kordig distinguishes two versions of Quine's thesis: (i) No hypothesis can be irrevocably falsified. (ii) No hypothesis can be immune to revision. To show the pragmatical self-referential inconsistency of both versions, Kordig argues as follows:

(i) The Quine-Duhem thesis is itself an hypothesis. By its own claim, it cannot be irrevocably falsified. Like the thesis itself, the negation of the Quine-Duhem thesis is an hypothesis which, according to the thesis, cannot be irrevocably falsified. Hence the denial of the thesis cannot be rejected with finality: It is possible to sustain the negation of the thesis, viz., that some hypothesis can be irrevocably falsified. Consequently, from the Quine-Duhem thesis, its falsification can be deduced. It is a self-falsifying pragmatical self-referential inconsistency, hence is not tenable.

Alternatively, (ii) the Quine-Duhem thesis is an hypothesis which claims that no hypothesis can be immune to revision. Hence it is open to revision. To revise an hypothesis, in Quine's view, is to change its truth-value. In other words, from the assumption that the Quine-Duhem thesis is true it follows that it may be false, in which case some hypothesis can be immune to revision. But this latter claim is in direct conflict with the original thesis. Once again, from the assumption that the Quine-Duhem thesis is true, it follows that it is false.

These three examples of pragmatical self-referential argumentation make two things clear: The claim that a position is pragmatically self-referentially inconsistent is forced, first, to suppose that the position attacked will acknowledge the legitimacy of its self-application. This is often problematic. As Passmore has ob-

served in connection with pragmatical self-refutation, the propounder of a position under criticism is always free in principle, "even if sometimes with almost inconceivable hardihood"[29] to deny the intentions attributed to him.

Secondly, provided the self-application of a position *is* accepted as legitimate by its propounder, it follows from a valid self-refutation that the statement of the position in question is self-falsifying. But it does not necessarily follow that what is claimed by the position cannot be the case. It may not be possible coherently to *state* the claim that there is no objectivity in science, or that there exists radical meaning variance, or that all hypotheses are open to revision, yet, it can be argued, it does not follow that any one of these cannot nevertheless be true. They may be true, but *this* possibility cannot be expressed consistently. The sceptical metaclaim, in attempting to say what cannot consistently be said, is doomed to self-referential inconsistency. The suspicion may linger that Feyerabend, Hanson, Hesse, Kuhn, Polyani, Quine, Smart, Toulmin may be right, but the suspicion cannot consistently be voiced. Among other things, this is what it means to say that a position is untenable.

Kordig's self-referential analyses do not, in their current formulation, focus on invariant conditions of discourse (although elsewhere there are some hints that he may eventually move in this direction[30]). His analyses appear to express self-referential *ad hominem* arguments in Johnstone's sense. That this is so appears to be confirmed by the vulnerability of Kordig's arguments to objections regarding the legitimacy of forcing the self-application of a position. (Objections of the second kind, "Even if the position *is* self-referentially inconsistent, it still may be true," are effectively silenced.)

Most self-referential analyses in philosophy of science have been pragmatical in focus, and have treated theories developed in philosophy of science *about* scientific theories.

29. Passmore (1961), p. 63.
30. In connection with Kordig's arguments that objectivity and meaning invariance *are* possible, see Kordig (1971a), (1971b), (1971c), (1973).

In contrast to such a pragmatical analysis of theories of theories, we turn to a metalogical argument concerning a particular scientific theory. Before doing so, let us recapitulate.

We recall that in a metalogical analysis of preconditions of referring an attempt is made to identify constraints which cannot be violated without projective self-referential inconsistency. A projective claim is not, like a pragmatical self-referential inconsistency, self-falsifying, but is *self-undermining*: A concept or proposition is used in a position in such a way that, literally and logically, precludes that the forms of reference involved can possibly obtain. A projective self-referential inconsistency results if one attempts to refer to an object o in such a way that denies one or more conditions which must be satisfied in order for it to be possible to refer to o at all. A self-undermining claim does not falsify itself, but is such that it is incoherent to associate any meaningful truth-value with the claim. In a somewhat metaphorical sense, pragmatical self-referential inconsistencies express factual short-circuits which involve either the intentions acknowledged by a position or certain invariant conditions of discourse, and which result in a falsification of that position. Projective self-referential inconsistencies express transcendental short-circuits which (a) involve self-validating preconditions of referring, and which (b) undermine the possible meaningfulness of a claim endorsed. These varieties of inconsistency, among others, represent ways in which conceptual structures may become dysfunctional and self-defeating.

VII

There has been strong opposition among philosophers to the Copenhagen interpretation of quantum mechanics. Among physicists, however, this interpretation has been the substructure for progress in theoretical and experimental research in microphysics for several decades. Contrary to this trend in physics, a bias in favor of realism and physical determinism was expressed in the opposing hidden variable interpretation of quantum mechanics. Numerous philosophers and a few physicists have claimed, in spite of the uncertainty

relations, that a microparticle in fact has a well-defined simultaneous position and momentum. From the standpoint of current quantum statistical mechanics, such a claim involves a metalogical self-referential inconsistency.

The uncertainty principle grows out of a calculus of operators. Two observables are said to commute if the observations are non-interfering. Quantum mechanics, specifically, matrix mechanics, asserts that for a class of dynamical variables, if P and Q are non-commuting operators, then P and Q are canonically conjugate quantities: that is, if a physical system is in a state for which P is determined with an accuracy ϵ, then there is a maximum limit to which Q may be determined, viz., $\eta = h/2\pi : \epsilon$.

The relation between such canonically conjugated variables is essentially one of uncertainty. Heisenberg's uncertainty principle, which expresses the logic of such variables, is normally discussed in the context of the noncommuting observables, position and momentum. However, there are analogous uncertainty relations involving other dynamical variables which cannot be precisely measured simultaneously, for example, energy/time, number/phase, etc.

The Copenhagen interpretation of quantum mechanics accepted the limitative results expressed in Heisenberg's application of the formalism of noncommuting matrices. In this view and in related formulations, the uncertainty relations do not merely represent technical limitations, but they rather constrain, in principle, what may meaningfully be stated in matrix mechanics, in wave mechanics, or in the more general so-called transformation theory. The micro-physical theory built on this foundation has been vigorously opposed by many philosophers and not very many physicists, among the latter Einstein, De Broglie, Jeffries, and Bohm. Of the arguments proposed, perhaps Bohm's is the only one which has not reduced simply to an endorsement of prejudices in favor of realism and complete physical determinism. Although it is not possible to go into the details of his view here, we may note that Bohm's rejection of the postulate of uncertainty did not evolve into more than a hopeful sketch of an alternative microphysical theory, one which

has received a sceptical response from physicists.[31] In discussing his alternative theory, Bohm speculated that

> ... the coordinates and momenta of individual atoms are hidden variables, which in a large scale system manifest themselves only as a statistical averages. Perhaps then, our present quantum mechanical averages are simply a manifestation of hidden variables, which have not, however, yet been detected directly.[32]

To show that this point of view expressed by Bohm and others is metalogically projective in relation to contemporary Copenhagen-based quantum theory, it is necessary to demonstrate these things:

- that the uncertainty relations have a presuppositional role in modern quantum statistical mechanics;
- that a denial of the postulate of uncertainty entails a denial of preconditions which must be satisfied in order for physical reference to specified dynamical variables to be possible.

It is rather straightforward to establish the first of these: Perhaps the most general assumption of existing quantum theory, as acknowledged even by Bohm, is that the state of a physical system "is completely specified by a wave function that determines only the probabilities of actual results that can be obtained in a statistical ensemble of similar experiments".[33] From this assumption, Bohm goes on to say,

> ... the uncertainty principle is readily deduced ... [It becomes a contradiction in terms to ask for a state in which momentum and position are simultaneously and precisely defined .... The uncertainty principle is ... a necessary consequence of the assumption that the wave function and its probability interpretation provide the most complete possible specification of the state of an individual system ....[34]

According to this view, the uncertainty relations can be derived from the assumption that the probability interpretation of the wave function constitutes a complete microphysical description. From the perspective of opponents to the Copenhagen interpretation, to claim, on this basis, that the postulate of uncertainty plays a presuppositional role would be to beg the question. It is precisely the foregoing assumption from which the principle of uncertainty is derived which they wish to question.

---

31. Heisenberg, Oppenheimer, Dirac, and Bethe expressed their strongest doubts concerning Bohm's proposal (in personal communications with Norwood Russell Hanson). Cf. Hanson (1958), p. 174.
32. Bohm (1952), p. 166.
33. Bohm (1952), p. 166.
34. Bohm (1952), p. 167.

Fortunately for our purpose, the reverse has also been shown: that solely from an operationally-based statement of the uncertainty relations the rest of quantum mechanics can be derived. In his famous proof, Von Neumann demonstrated[35] that, indeed, the uncertainty relations make up, as Hanson put it, "the logical backbone of *all* quantum theory."[36]

Two further remarks may exhibit some of the force behind this demonstration. First, so-called "interference terms" occur in quantum mechanics. They are not understood simply as products of probabilities, but are functionally defined as products of Ψ functions. Put somewhat differently, the noncommuting nature of such dynamical parameters as position and momentum is entailed by the nature of the Ψ function.

Secondly, it is interesting to note that, as a consequence, the algebraic analog of a statement simultaneously specifying precisely de-

35. Von Neumann (1955), Chapters IV, VI, especially pp. 323ff.

36. Hanson (1967), p. 46.

It should be noted that Bohm did not disagree with von Neumann's argument. Bohm conceded that as long as the usual rules of calculating quantum-mechanical probabilities are in force, it is inconsistent to postulate a set of hidden parameters which simultaneously determines the results of measurements of noncommuting observables. (Bohm (1952), II, p. 187.) Bohm's proposal essentially sought to modify these rules: in particular, to consider such observables as position and momentum as "potentialitieswhose precise development depends just as much on the observing apparatus as on the observed system. In fact, when we measure the momentum "observable", the final result is determined by hidden parameters in the momentum-measuring device as well as by hidden parameters in the observed electron. Similarly, when we measure the position "observable", the final result is determined in part by hidden parameters in the position-measuring device." (*Ibid.*) Bohm's proposal acknowledged that these two measurements are mutually exclusive since they depend on "mutually exclusive arrangements of matter that must be used in making different kinds of measurements." (*Ibid.*, pp. 187-188.)

In spite of this block to simultaneous measurements of, e.g., position and momentum, Bohm wished to be able to claim that both observables are in reality precisely determined in a physical system. To maintain this, Bohm describes the preparation of a physical system in a state "in which the Ψ-field and the initial particle position and momentum are precisely known." (*Ibid.*, p. 185.) According to Bohm's theory, then, it is possible to measure only one of these observables precisely: it is necessary to *infer* the value of the other on the basis of formal relations of the theory.

Bohm wished to preserve precise simultaneous determinability of both observables, not merely by inference, but in fact. The realistic claim, in the context of his own theory, is projective: The microphysical claim that both parameters are precisely defined and physically real presupposes that in principle both can simultaneously be measured. Yet, as we have seen, Bohm's position accepts the constraint that measurements of position and momentum are mutually exclusive.

fined values for position and momentum itself is without meaning in quantum statistical mechanics. The absence of meaning here is due to conflict with the rules of formation and transformation employed in the formalism. But there is another, perhaps more compelling, reason for its meaninglessness:

As long as an alternative, comparably detailed microphysical theory is unavailable, the physical meaningfulness of a microphysical claim – e.g., relating to mutually interfering observables – will be understood in terms of prevailing quantum statistical theory. The uncertainty relations have the status of presuppositions – conceived of as rule-based constraints – within the conceptual structure of the theory. The uncertainty relations are nothing more than the expression of a limitative postulate required in a calculus of operators. Now, a hidden variable theorist wishes to refer to subatomic events as currently understood in the context of existing quantum theory. He wishes, furthermore, to claim that mutually interfering observables actually posses well-defined simultaneous values. Such a claim is clearly projective: The hidden variable theorist refers to a pair of observables which are *essentially* defined in a noncommuting sense, and in so doing explicitly denies a condition which logically is forced on our current understanding of interfering observables. The condition he denies is a precondition which must be satisfied in order for it to be possible for him, or anyone else, to refer meaningfully in the theoretical context in question to such observables. It is not that what the hidden variable theorist says is self-falsifying; rather, his claim is self-undermining in terms of its possible meaningfulness.

Should an alternative microphysics someday be developed as Bohm hoped, in terms of which microparticles meaningfully may be said simultaneously to possess precisely determined positions and momenta, time and energy, number and phase, etc., the above conclusion will stand unaffected. The uncertainty relations essential to Copenhagen quantum mechanics remain essential in physics as long as that theory is held. A second theory in which this is not the case refers, in a quite literal and logical sense, to objects which are defined in an essentially distinct way. A physical metatheory – which

correlates predictions made by Copenhagen quantum mechanics and a possible alternative Bohm microphysics – would enable physicits to evaluate the comparative usefulness of the two theories. The predictive value of the competing theory conceivably might be greater than that of the Copenhagen view, in which case it would have to give way to the new theory. Thus, where Bohm's hidden variable claim expressed in its present conceptual environment is projective, a corresponding claim asserted in the context of a fully developed, alternative microphysics, is trivial. The two claims can by no means be reduced to the same claim: One is self-undermining, while the other is best likened to a tautology.

Steven J. Bartlett is Professor of Philosophy at Saint Louis University. He attended the University of Santa Clara, Raymond College of the University of the Pacific, the University of California at Santa Barbara, and received his Ph.D. from the Université de Paris, where his research was directed by Paul Ricoeur. He has been a fellow in philosophy and mathematics at the Center for the Study of Democratic Institutions, a visiting research fellow at the Max Planck-Institut in Starnberg, West Germany, and has taught at the University of Florida, the University of Hartford, and Saint Louis University. Dr. Bartlett is author of a book on mathematical logic, a monograph on general problem-solving and many articles.

REFERENCES

Bartlett, Steven J., *A Relativistic Theory of Phenomenological Constitution: A Self-Referential, Transcendental Approach to Conceptual Pathology*, 1970, *Diss. Abs. Internatl.* No. 7905583.

Bartlett, Steven J., "Phenomenology of the Implicit", *Dialectica: Revue internationale de philosophie de la connaissance*, Vol. 29, No. 2-3 (1975), pp. 173-188.

Bartlett, Steven J., "The Idea of a Metalogic of Reference", *Methodology and Science*, Vol. 9, No. 3 (1976), pp. 85-92.

Bartlett, Steven J., "A Metatheoretical Basis for Interpretations of Problem Solving Behavior", *Methodology and Science*, Vol. 11, No. 2 (1978), pp. 59-85.

Bohm, David, "A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables", *Physical Review* LXXXV (1952); Part I, pp. 166-179; Part II, pp. 180-193.

Collingwood, R.G., *The Idea of History* (Oxford: Oxford University Press 1946).

Donagan, Alan, *The Later Philosophy of R.G. Collingwood* (Oxford: Oxford University Press 1962).

Feyerabend, Paul K., "Explanation, Reduction, and Empiricism", *Minnesota Studies in the Philosophy of Science*, Vol III, ed. H. Feigl and G. Maxwell, (Minneapolis: University of Minnesota Press 1962), pp. 28-97.

Fitch, Frederic Brenton, "Self-Reference in Philosophy", in Fitch (1952), pp. 217-225. (First published in *Mind*, Vol. 55 (1946), pp. 64-73.)

Fitch, Frederic Brenton, *Symbolic Logic* (New York: The Ronald Press 1952).

Fitch, Frederic Benton, "Universal Metalanguages for Philosophy", *Review of Metaphysics* 17 (1963-64), pp. 396-402.

Hanson, Norwood Russell, *Patterns of Discovery* (Cambridge: The University Press 1958).
Hanson, Norwood Russell, "Philosophical Implications of Quantum Mechanics", *Encyclopedia of Philosophy*, Vol. 7 (New York: Macmillan 1967), pp. 41-49.
Hesse, Mary, "A New Look at Explanation", *Review of Metaphysics* 17 (1963), pp. 98-108.
Hesse, Mary, "Fine's Criteria of Meaning Change", *The Journal of Philosophy* 65 (1968), pp. 46-52.
Johnstone, Henry, W., Jr., *Philosophy and Argument* (University Park, Penn.: Pensylvania State University Press 1959).
Johnstone, Henry W., Jr., "Argumentation and Inconsistency", *Revue internationale de philosophie* 15, No. 58 (1961), pp. 353-365.
Johnstone, Henry W., Jr., "Self-Refutation and Validity", *The Monist* XLVIII (1964), pp. 467-485.
Kordig, Carl R., "Objectivity, Scientific Change, and Self-Reference", *Boston Studies in the Philosophy of Science*, Vol VIII (Dordrecht, Holland: D. Reidel Publishing Co. 1970), pp. 519-523.
Kordig, Carl R., "The Theory-ladenness of Observation", *Review of Metaphysics* 24 (1971,a), pp. 448-484.
Kordig, Carl R., "Scientific Transitions, Meaning Invariance, and Derivability", *Southern Journal of Philosophy* 9 (1971,b), pp. 119-125.
Kordig, Carl R., "The Comparability of Scientific Theories", *Philosophy of Science* 38 (1971,c), pp. 467-485.
Kordig, Carl R., "Observational Invariance", *Philosophy of Science* 40 (1973), pp. 558-569.
Kuhn, Thomas S., *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press 1962).
Lorenzen, Paul, *Normative Logic and Ethics* (Mannheim/Zürich: Bibliographisches Institut 1969,a).
Lorenzen, Paul, *Einführung in die operative Logik und Mathematik* (Berlin: Springer 1969,b).
Passmore, John, *Philosophical Reasoning* (London: Gerald Duckworth 1961).
Quine, Willard V., "Two Dogmas of Empiricism", in *From a Logical Point of View*, Second Edition (New York: Harper and Row 1963), pp. 20-46.
Quine, Willard V., *Methods of Logic* (New York: Holt, Rinehart and Winston 1972).
Russell, Bertrand, *Our Knowledge of the External World* (London: George Allen and Unwin 1972 – first published 1914).
Rynin, David, "Donagan on Collingwood: Absolute Presuppositions, Truth, and Metaphysics", *Review of Metaphysics* XVIII (1964), pp. 301-333.
Scheffler, Israel, *Science and Subjectivity* (New York: Bobbs-Merrill 1965).
Smart, J.J.C., "Conflicting Views about Explanation", *Boston Studies in the Philosophy of Science*, Vol. I, ed. R.S. Cohen and M.W. Wartofsky (New York: Harper and Row 1953), pp. 157-169.
Toulmin, Steven, *Foresight and Understanding* (New York: Harper and Row 1961).
Von Neumann, John, *Mathematical Foundations of Quantum Mechanics* (Princeton, N.J.: Princeton University Press 1955); originally published as *Mathematische Grundlagen der Quantenmechanik* (Berlin: Verlag Julius Springer 1932).

CHAPTER 12

Referential Consistency as a Criterion of Meaning

STEVEN J. BARTLETT

REFERENTIAL CONSISTENCY AS A CRITERION OF MEANING*

Criteria of meaning which have been proposed in the past have failed to persuade general acceptance. They have usually endorsed then-current scientific practice, or have favored the adoption of a special, usually empirical, framework. The historical failure of criteria of meaning has been due to their apparently arbitrary status as standards external to the sets of statements to which they would apply. Often, such criteria have also failed to qualify as meaningful in the test of self-application.

It is my purpose here to show that there is available to us a criterion of meaning which must be satisfied in order for individual claims, concepts, and frameworks to qualify as "meaningful". The criterion I shall recommend is that of "referential consistency". It is proposed as a criterion of meaning in the largely negative sense that non-satisfaction of the criterion involves a certain type of meaninglessness that has received little attention. The criterion developed here therefore does not express a sufficient condition of meaningfulness. One may indeed seriously doubt whether a sufficient condition can be formulated. As a result of this limitation of focus, little will be found here about the nature of meaning. On the other hand, the criterion proposed defines an important *lower limit* of meaning, below which claims, concepts, and frameworks become self-undermining. It is in this latter sense that the criterion proposed can provide a useful tool for internal analysis and criticism.[1]

The criterion I shall suggest has these rather unique properties: Acceptance of the criterion is non-arbitrary or compelling in a sense we shall explore briefly. And applications of the criterion avoid begging the question in a way in which appeals to external standards do not.

Logical criteria for evaluating, e.g, the validity of an argument, or for assessing the consistency of a theory, define for us the limits of acceptability which argumentation or theory construction endorses. To a large degree, such criteria are arbitrary in the sense that they can be changed if our purposes are served by such a change. Seen as conventions we accept in the light of our objectives,[2] the criteria

which delimit what we will accept are seldom, if ever, absolute. That is, we are not normally compelled, on pain of incoherence, to accept certain particular criteria rather than certain others, although it is often the case that, if we are to hold to our purposes, we must abide by these or related criteria if we are to accomplish what we intend.[3] In general, then, I shall call a criterion *non-arbitrary* or *compelling* if non-satisfaction of that criterion precludes achieving the task at hand. We shall look at this claim in more detail in a moment.

The criteria which define what we mean, e.g., by 'validity' and 'consistency' are "logically arbitrary" in several ways. If we detect that a criterion, or equivalently here, a rule, has been broken, we are free to amend the rule (and perhaps in so doing change the ends which the rule may serve), or correct the violation, or leave things as they are, or shift our perspective, perhaps to a more general point of view, and perceive the breaking of the rule as conforming to a more general rule in relation to which it is no longer identified as a violation. And we may have other options. But whatever the special nature of the case may be, criteria of the sort used to assess the validity of arguments and the consistency of theories constitute logically arbitrary rules for playing certain games: rules are the logical features of practical activity; the control which they make possible is a control which we choose to have, and we are free to choose otherwise.

In relation to our chosen purposes, then, logical criteria seldom compel us by reason of logic alone to accept these criteria and no others. There is, often and in general, a sense of "open-texture" about our objectives. The formal constraints we do accept may be selected because they reinforce other ends we intend: economy, comprehension, concinnity, etc. *How* we do or should make selections from among alternative, logically arbitrary criteria will not be examined here.

From the standpoint of the criteria we accept, our purposes are underdetermined or specified with a degree of vagueness to just the degree that these criteria are logically arbitrary. It is perhaps fair to say that attempts to delimit meaning by means of a necessary and sufficient criterion have failed because of this logical arbitrariness. The numerous criteria that have been recommended for detecting meaningless concepts and statements have very much the same status as do criteria which permit evaluations of validity, consistency, etc.

Criteria of meaning have come to be considered in the same game-relative light as have rules of logical evaluation.

For example, Hume, Schlick, Ayer, and Carnap have proposed these as criteria of meaning:

For Hume: expression of abstract or empirical reasoning.[4]

For Schlick: association of conditions with a proposition or question which define what experience(s) would make that proposition true, or which would if satisfied answer that question.[5]

For Ayer: verifiability, reflecting an individual's knowing how to verify a proposition which is factually significant to him.[6]

For Carnap: ability to give rules according to which observable effects can be deduced,[7] or alternately, expression of factual content.[8]

These criteria, not exhaustive of those proposed in the literature, nor yet mutually exclusive, share two characteristics: First, from a non-partisan viewpoint, it may be fair to say that acceptance of one or more of these criteria is a function of one's purposes. Second, neither Hume, nor Schlick, nor Ayer, nor Carnap, nor any other proponent of a criterion of meaning apparently has been able to show that acceptance of a certain criterion of meaning compels assent, i.e., is non-arbitrary in the sense we have sketched.

This observation would not reflect a negative judgment if, as could be claimed, we wish a criterion of meaning to function with the same measure of arbitrariness in the framework of a set of concerns as does a rule-based convention of logical evaluation.[9] But this state of affairs would clearly not satisfy authors of meaning criteria.

Criteria of meaning, then, have functioned in an *external* capacity: When they are applied, they are used to evaluate statements, concepts, or frameworks, as it were, from the outside. Criteria of meaning, understood as stipulative, normative conventions, can only be recommended in a manner which seeks to persuade our acceptance, since they do not, in and of themselves, compel assent.[10]

One of the most persuasive cases that can be made on behalf of the choice of a certain criterion of meaning is that its meaningfulness follows from its self-application.[11] If the criterion recommends that meaning be identified with expression of factual content, for example, it may be argued that 'factual content', understood in terms of operations which define the criterion, itself expresses factual content.

However, the self-applicability of a criterion of meaning, when

assured, at most insulates the use of the criterion from internal inconsistency, and may strengthen the *feeling* that its choice is not totally arbitrary. Beyond this, self-applicability does not do much: The decision to adopt a particular criterion of meaning remains external to the class(es) of statements and concepts to which it is to apply.

REFERENTIAL CONSISTENCY AS AN INTRINSICALLY DETERMINED CRITERION OF MEANING

In the view I have attempted to represent, rules for evaluating logical validity and consistency and criteria of meaning share the property of arbitrariness as game-relative conventions. The selection of such rules and criteria hence may be considered predominantly to be a function of our practical concerns. With respect to the decision to adopt a particular criterion, there is little that can be said if more than practical justification is desired. In a given field of study, rule-based evaluative conventions of one kind or another may be convenient, expedient, or necessary in practice. If one chooses to work in that field, he may have need of some externally imposed evaluative conventions. But the use of such external standards of evaluation cannot, as we have seen, be expected to be non-arbitrary and compelling.[12]

Fortunately, *there does exist a logically compelling basis for evaluation, a basis which one cannot not accept.* I have called this basis for evaluation 'referential consistency'.[13]

Referential consistency does not represent an externally imposed convention, a normative stipulation, an arbitrarily endorsed special rule or criterion. The approach to referential consistency described here rather has the character of a *metalogic*, in terms of which "preconditions of reference" in special contexts can be studied. In rough terms, initially, referential consistency is a metalogical criterion or rule of evaluation which addresses, *intrinsically*, the context-relative use of expressions, statements, or concepts. A special set of evaluative rules or criteria is not applied across the board in an external way, but rather *attention is given to those conditions which must be satisfied in a given context in order for references made in that context to be possible at all.* The results of applying such a metalogical criterion of referential consistency are non-arbitrary, both

because a special criterion is not imposed externally, and because these results compel assent – one cannot reject them in a given context of reference.

A short account of the proposed metalogic of reference will be given here. A complete formulation of the general theory will be found elsewhere,[14] as are illustrations of certain applications of the metalogic.[15]

A METALOGIC OF REFERENCE

For the sake of simplicity, I limit my treatment here to the set of *referring sentences* (alternatively, *propositions*) $\mathscr{P} = \{p_1, p_2, \ldots, p_n\}$ where p_i may refer to any one or more o_i of a set of objects of reference $\mathcal{O} = \{o_1, o_2, \ldots, o_n\}$, and may possess any truth-value of a set of possible values $V = \{0, 1, \ldots, n\}$, where $n \geq 3$.[16] By the 'significant range of V' is meant '$\{0, 1, \ldots, n-1\}$'. (A discussion of the value v_n follows below.) It is clear that the significant range of V is bivalent when $n - 1 = 1$, with '0' and '1' representing the values 'false' and 'true', respectively.

Some definitions are called for.

(D1) A *particular* is a possible object of identifying reference.

Alternatively,

(D2) An *identifying reference* is such that an ascription to that which can be the subject of an ascription (namely, a particular) establishes that what is ascribed (one or more properties, relations, a description, etc.) and that that to which the ascription is made are one and the same (identification).

D1 contains the there undefined concept of identifying reference, while D2 leaves undefined the concepts of particular, description, property, relation, identification, and ascription.

In the interests of simplicity we will retain D1, permitting the concept of identifying reference to be primitive. However, it may be useful to introduce an interpretation concerning the use of 'identifying reference'.

In what follows, 'R' is used to express a ternary relation between a person, whose proper name may be assumed as a value by a variable

'α' ranging over a set of proper names for persons, and a space–time coordinate which is a value of a variable 'σ' taking as its values specific space–time coordinates. When identifying reference (hereafter simply called 'reference') to an object obtains, Rmo_is uniquely determines o_i in relation to a person m, at a certain space–time coordinate s:

$$(1) \qquad (x)(Rmxs \;\&\; \cdot\, x \in \{o_1, o_2, \ldots, o_n\} :\supset$$
$$-(\exists y)(Rmys \;\&\; \cdot\, y \in \{o_1, o_2, \ldots, o_n\} :\&\; x \neq y).^{17}$$

From this point of view, the concept of reference is used to address the metalogical properties of identification; that is to say, possession of an identity is presupposed in connection with any particular, and all particulars are possible objects of reference, i.e., can be identified.[18] (It is important, then, to observe that the term 'reference' is *not* used in a way that entails the existence of psychological processes, intentions, etc., although these dimensions of referring need not be excluded if we wish to talk about them.)

Let $p_i \supset R\alpha o_i\sigma$ express the claim that the use by a person α at a space–time position σ of a referring sentence p_i entails reference to an o_i, if p_i has a value in the significant range; in other words, $R\alpha o_i\alpha$ follows from p_i whether the value of p_i is T or F.[19] The claim that is implicit here is that referring sentences of $\mathscr{P}$ are such that reference obtains to some o_i provided only that the p_i of $\mathscr{P}$ have truth-values in the significant range: hence, even when a p_i = F, reference is considered to obtain to some o_i which can serve to justify the claim to the effect that p_i = F.

A p_i is said to be *self-referentially inconsistent* in three cases which we distinguish here. (1) When $p_i \supset R\alpha o_i\sigma$ and $o_i = p_i$, then p_i exhibits *sentential* or *propositional self-reference*, depending upon whether p_i is considered as a sentence or as the expression of a proposition. If p_i is self-referential in either of these two ways and p_i claims of itself that it is false, then, when V is bivalent, p_i is true iff it is false. Such a p_i comprises a *paradox-generating* self-referential inconsistency. Many of the semantical paradoxes are clearly of this form.

(2) When $p_i \supset R\alpha o_i\sigma$ and $o_i = P_{p_i}$, where 'P' designates a pragmatic (or performatory) aspect of the use made of p_i by α at space–time position σ, then p_i is termed *pragmatically* (or performatively) *self-referential*. If p_i is pragmatically self-referential and p_i is such that if p_i is asserted or otherwise is used in a manner such that P_{p_i} falsifies

p_i, then, when V is bivalent, p_i is said to be *self-refuting*. The assertion, for example, "This assertion does not refer to an x such that Fx", for interpretations of 'x' and 'F', expresses a self-refuting self-referential inconsistency. Ramsey's familiar example, "I can't say 'cake'", when uttered by anyone, accordingly may be seen to be self-refuting.

(3) When $p_i \supset R\alpha o_i\sigma$ and $R\alpha o_i\sigma \supset R\alpha M_{p_i}\sigma$,[20] where '$M_{p_i}$' designates a "precondition of reference" which must be satisfied in order for p_i to have a value in the significant range, then p_i is termed *metalogically self-referential*. If p_i is metalogically self-referential and p_i is such that p_i denies one or more conditions which must be satisfied in order for it to be possible to assert, or otherwise use, p_i significantly, then p_i is said to be *projective*, or $\tilde{p}_i$.[21]

The expression 'precondition of reference' is associated with the following equivalent senses: 'M_{p_i}' designates a "precondition of reference" if, in order for reference to be possible in a particular context of reference, M_{p_i} must be satisfied; M_{p_i} is a necessary condition of possible reference; M_{p_i} qualifies as a "precondition of reference" iff it designates a condition the non-satisfaction of which in a particular context of reference results in projection.

When V is bivalent, a metalogically self-referentially inconsistent p_i makes, with a putative value T or F, an ascription a of some object of reference o_i. If $p_i = \mathrm{T}$, then a applies to o_i, or $a(o_i)$; if $p_i = \mathrm{F}$, then $-a(o_i)$. In either case, possible reference to o_i is presupposed.

(2) $\qquad a(o_i) \vee -a(o_i). \supset \Diamond R\alpha o_i\sigma$

In short, when V is bivalent,

(3) $\qquad \tilde{p}_i \equiv a(o_i) \vee -a(o_i). \;\&\; -\Diamond R\alpha o_i\sigma,$

where $-\Diamond R\alpha o_i\sigma$ is implied by the projective denial of one of the conditions which must be satisfied in order for it to be possible significantly to assert p_i.

$$\vdash p_{i(\mathrm{T}\vee\mathrm{F})},\ p_i \supset R\alpha o_i\sigma,\ R\alpha o_i\sigma \supset \Diamond R\alpha o_i\sigma,\ \Diamond R\alpha o_i\sigma \supset M_{p_i},$$

(4) $\qquad p_i \supset -M_{p_i} \Vdash -\Diamond R\alpha o_i\sigma$

The self-referential inconsistency of a projection is rendered explicit when the consequent of (2) and the conclusion of (4) are conjoined.

P. W. Bridgman's hypothesis to the effect that the entire physical universe is shrinking homogeneously, i.e., in a manner such that all

operations of measurement are correspondingly affected, may be seen to be projective. In order for the hypothesis to be significant in a bivalent system, in order for reference to be made to "universal homogeneous shrinkage", Bridgman argues that it must be presupposed possible to detect relevant changes in relative size of the physical universe. This is essential to the meaning of the concept of shrinkage. However, by hypothesis, universal homogeneous shrinkage rules out that the precondition of reference, possible detection of the alleged change in relative size, can be satisfied. Hence the hypothesis is projective.[22]

In an intuitive sense, $p_i \supset \vec{p}_i$ will hold when p_i conflicts self-referentially with preconditions which must be granted in order for the value of p_i to fall in the significant range. A projective assertion consequently involves a special kind of self-referential inconsistency. Our main interest here is in projective forms of reference.

For a bivalent range of significance, $p_i = \mathrm{T}$ when

$$\text{`}p_i\text{'} \text{ is true iff } p_i \quad \text{(Tarski's definition);}$$

and $p_i = \mathrm{F}$ when

$$\text{`}p_i\text{'} \text{ is false iff } -p_i.$$

When p_i is projective, p_i is said to have value μ

$$\text{`}p_i\text{'} \text{ has value } \mu \text{ iff } \vec{p}_i.$$

Here, 'μ' represents the value 'projective meaninglessness' which lies outside the significant range of values $\{0, 1, \ldots, n-1\}$. It should be clear from the nature of a projective assertion that its value cannot be identified with any of the values in its putative significant range since one or more conditions are denied which must be satisfied in order for p_i to have *any* value in the significant range. The self-referential inconsistency of a projective assertion is of a kind which literally and logically precludes that the assertion *can* possess a value in the significant range. In some contexts there may be some latitude of choice whether to consider an assertion to be meaningless or false: e.g., in the case of the infamous 'The present king of France is bald'. From the standpoint of metalogic of reference, however, no other option is available: The value of a projective assertion must fall outside the significant range, hence it is appropriate to identify its value μ with meaninglessness.

A p_i is said to be *self-validating* in the case where $-p_i$ is metalogically self-referentially inconsistent. Conversely, p_i is said to be metalogically self-referentially inconsistent in the case where $-p_i$ is self-validating. I.e.,

(5) $(x)(x \in \mathscr{P} \cdot \& \, Fx :\supset \cdot G - x)$ and

$(x)(x \in \mathscr{P} \cdot \& \, Gx :\supset \cdot F - x)$,

where F is the property '... is self-validating' and G is the property '... is self-referentially inconsistent'.

It follows that for any $p_i \supset \vec{p}_i$, and hence when p_i has value μ, the equivalent claims 'the value of p_i does not fall in the significant range', 'p_i is not significant', 'p_i is meaningless' self-validate since the denial of any one entails self-referential inconsistency. *For this reason, referential consistency, as a metalogical criterion of meaning, cannot not be accepted.* Referential consistency is, in other words, a self-validating criterion which must be satisfied in order for claims to be meaningful.

It may be noted that the significant range of the set V of possible values of a p_i has been left unspecified, although in general we have defined the significant range to coincide with $\{0, 1, \ldots, n-1\}$. Leaving the significant range unspecified in this way has the advantage of flexibility, since, in some contexts of reference, we may wish to be able to assign values representing indeterminacy, statistical probabilities, etc., to a p_i (for example, in quantum logics). Although no decision has been made, then, in favor of bivalence in V, the following metametalanguage formulation is implied by the principle of bivalence, without implying it:

(i) Every referring sentence of $\mathscr{P}$ either has a value in the significant range, or it does not.

Adoption of this metalogical version of the principle of bivalence entails that all metalogical statements assigning values from $\{0, 1, \ldots, n-1, n\}$ – (from the range of possible values from falsity (0) to a designated value (1 in a bivalent system) to μ) – to a p_i are themselves true or they are not. In fact, (i) entails

(ii) There exist in principle possible procedures which yield a yes or no determination for any metalogical value-assigning statement about members of $\mathscr{P}$.

It will be evident to the reader that the assertion of (i) conjoined with the rejection of (ii) constitutes a projective assertion. Consequently, we shall regard (ii) as entailed, in a self-validating manner, by (i).[23]

By way of illustration, let us assume V is bivalent; hence the significant range comprises values T (1) and F (0) with μ representing the value of projective assertions. In effect, then, the set of sentences or propositions $\mathcal{P}' = \{p_1, p_2, \ldots, p_n\}$[24] will be, for the purposes of assessing referential consistency, three-valued within a bivalent metalanguage. (Such a three-valued representation can be reduced, as we shall see, to a two-valued representation, with T, F = Ψ, where 'Ψ' simply indicates a value in the significant range.)

Matrices for conjunction and negation suitably take the form proposed by Bochvar:[25]

–	
F	T
T	F
μ	μ

&	T	F	μ
T	T	F	μ
F	F	F	μ
μ	μ	μ	μ

Where μ is the value of a projective assertion, the above matrices make clear that the negation of a projection remains projective, while the conjunction of a projection with a significant assertion infects the compound statement, so to speak, with meaninglessness. The projective character of one conjunct may undermine the referential consistency of the other conjunct. The matrix for conjunction avoids this eventuality.

Other connectives are easily defined:

$A \vee B$ for $-(-A \,\&\, -B)$

$A \supset B$ for $-(A \,\&\, -B)$

$A \equiv B$ for $(A \supset B) \,\&\, (B \supset A)$,

so that the following matrices are determined:

$\vee$	T	F	μ
T	T	T	μ
F	T	F	μ
μ	μ	μ	μ

$\supset$	T	F	μ
T	T	F	μ
F	T	T	μ
μ	μ	μ	μ

$\equiv$	T	F	μ
T	T	F	μ
F	F	T	μ
μ	μ	μ	μ

From these matrices it can readily be seen that once part of an expression assumes the value μ, the expression automatically assumes the value μ. (The same rationale applies here as in the case of conjunction.) It is also evident that if all μ-rows and -columns are *eliminated*, the matrix is reduced to the normal two-valued one. If one sets $T, F = \Psi$, then it is clear that the elimination of statements of value μ leaves a set of statements having the value Ψ, statements which are in the significant range. This is an obviously desirable property of a necessary, not sufficient, criterion of meaning: its application will lead to the elimination of certain meaningless statements, leaving untouched all candidates which may be significant (and perhaps to which other necessary conditions of meaning may be applied.)

The metalogical criterion of meaning which emerges from this discussion is both *non-arbitrary* and *compelling*. It is *non-arbitrary* because the criterion is intrinsically informed by the special character of individual contexts of reference. It is *compelling* because one cannot at one and the same time consistently use expressions, sentences, or concepts referringly yet undermine their capacities to refer. Finally, a metalogical criterion of meaning which is defined in terms of referential consistency is *self-validating*: rejecting its application leads to projection.

In such a metalogical understanding of meaning, criteria for evaluating consistency and significance are determined as a function of one's needs and interests in making reference to certain kinds of objects. Within any specific context of reference, with these needs and interests in view, intrinsically determined criteria for evaluating internal consistency and significance *merge*, from the standpoint of a general metalogic of reference. They provide critical tools for appraising the meaningful use of expressions, sentences, or concepts in that context. Referential consistency is, in short, a contextually determined, yet non-arbitrary, compelling, and self-validating criterion of meaning.

In conclusion, it may be of interest to consider the relationship between a metalogical conception of meaning as a function of referential consistency, and the problem of putative meaningfulness.

THE PROBLEM OF PUTATIVE MEANINGFULNESS

Let p_i be a sentence or proposition in the context of a system SI which permits unambiguous identifying and re-identifying reference[26] to a set $\mathcal{O} = \{o_1, o_2, \ldots, o_n\}$ of objects. Let it further be agreed that a p_i is asserted to have a truth-value in the significant range, i.e., $\neq \mu$.

Upon analysis, it is determined that $p_i \supset \tilde{p}_i$, because $p_i \supset R\alpha o_i\sigma$, while $M_{pi} \supset -\Diamond R\alpha o_i\sigma$. From a metalogical frame of reference, M, then, we associate with p_i a truth-value μ not in the significant range. Note that *this* claim is an assertion about p_i-in-SI, and hence is a metalogical claim whose truth-value is determined on a bivalent metalogical basis.

It will be evident that the problem of putative meaningfulness is resolved. This problem has been pointed to by opponents to the use of meaning criteria. They have argued that, on the one hand, we have an expression, sentence, or concept which is used in various contexts, and in what is considered to be a meaningful fashion. Yet, upon application of a criterion of meaning, the alleged meaning is supposed to be given up, and the matter closed. The initially perceived meaning is not according to this view, "really meaningful". Such a suggestion runs counter to belief, i.e., is literally (not logically) paradoxical. This counter-intuitive character of results that stem from the application of meaning criteria decidedly has not promoted the popularity of criteria of meaning.

However, the quasi-paradoxical appearance of the problem of putative meaningfulness is simple to dispel: From the standpoint of SI, p_i is used to refer to an o_i so that o_i is uniquely determined. From the standpoint of M, reference is made to p_i-in-SI and reveals, through an analysis of p_i's referential preconditions, that the assertion of p_i-in-SI undermines p_i's capacity to refer to o_i.

If we associate a "meaning spectrum" V' with p_i such that $V' = \{0, 1, \ldots, n\}$, where $n = \mu$, then for any $0 \leq v_i < n$, v_i falls in the significant range V of V'. While the assignment of any v_i up to and including v_{n-1} may be made from the standpoint of SI, μ-assignments require recourse to a metalogical frame of reference M. In short, the possibility of detecting that a p_i has value μ is essentially a function of M's referential capacity. A metalogical statement S asserting that p_i is projective in SI, independently of M, itself is projective, as the reader may confirm.

There is, then, no problem with respect to putative meaningfulness from this viewpoint. What opponents to the use of meaning criteria very likely have in mind falls appropriately in this view under the heading of "making mistakes" and "detecting errors". When one makes a mistake without realizing it at the time, and later discovers his error, the passage of time provides what is, in effect, a metasystem which permits reference to what is retained in memory: From this vantage point, one compares what one remembers having thought earlier with what one now knows, and claims, in retrospect, that a mistake was made at the earlier time. The same may be said in the present case: The use of p_i to refer to o_i in SI was erroneous because p_i can be shown to be projective in M.[26] Hence, making an assertion which can be shown to be projective and hence meaningless in the sense developed, is simply to make one of many different kinds of possible mistakes.[27]

To remind us of this, it is convenient to view μ-assignments as involving, in a very literal sense, a *shift of significance.* Assumption of a metalogical frame of reference with respect to a projective assertion p_i results in a shift in p_i's putative truth-value (in SI) to μ (in M). Such a shift in significance is essentially a function of the metalogical frame of reference used. It is clear that a more comprehensive account of results proceeding from applications of a metalogic of reference would reveal many such *shifts to the value* μ of expressions, sentences, and concepts erroneously believed to be significant.[28]

Saint Louis University

NOTES

* Research reported here was supported in part by a grant from the Max-Planck-Gesellschaft.

[1] The general concern, to identify and eliminate meaningless statements and concepts from technical and/or ordinary discourse reflects a long tradition in which logic and philosophy together have sought to clarify our conceptual structure, and exhibit departures from sense. For example, Kant made mention of the need for a "negative science", a *phaenomenologia generalis*, which would undertake what might now be construed as a kind of "meaning-sorting", to insure that only meaningful propositions remain as the subject for subsequent analysis. (In a letter to Lambert, dated September 2, 1770.) The list of names in this tradition could be expanded almost indefinitely.

[2] For a statement of the view that logical rules essentially comprise conventions we

agree upon, and for additional references, see, e.g., Haskell B. Curry, *Outlines of a Formalist Philosophy of Mathematics* (Amsterdam: North-Holland 1957). See below, Note 12.

[3] Wittgenstein has given considerable attention to the relationship between using rules and achieving practical ends. See, e.g., his *Remarks on the Foundations of Mathematics*, edited by G. H. von Wright, R. Rhees, and G. E. M. Anscombe, trans. by G. E. M. Anscombe (Oxford: Basil Blackwell 1956): I – 9, 20, 131, 162; V – 31ff; and passim.

[4] Hume, *Enquiry*, sec. XII, iii.

[5] Moritz Schlick, 'Positivism and Realism', in A. J. Ayer (ed.), *Logical Positivism* (New York: Free Press 1959), pp. 82–107.

[6] A. J. Ayer, *Language, Truth and Logic* (London: Gollancz 1936; rev. ed. 1946), p. 35.

[7] Rudolf Carnap, *Philosophy and Logical Syntax* (London: Kegan Paul 1935), pp. 13–14.

[8] Carnap, *The Logical Structure of the World and Pseudoproblems in Philosophy*, trans. by R. George (Berkeley: University of California Press 1967), pp. 325ff and passim.

[9] There is certainly this sociological difference: Certain logical rules are "hard-programmed" in our culture, so that their *rejection* is counter-intuitive, as, for example, in the proposed use of non-distributive lattices in quantum theory. The matter is the other way around when it comes to criteria of meaning, since *violations* of the criteria heavily populate the domains of ordinary, and of some technical, discourse. And, to this extent, *acceptance* of certain of the proposed criteria of meaning frequently results in a counter-intuitive reaction in our culture.

[10] Carnap's introductory sentences in his *Logical Structure of the World* come to mind: "What is the purpose of a scientific book? It is meant to convince the reader of the validity of the thoughts which it presents."

[11] On the requirement that a theory of meaning be self-referentially meaningful, see R. J. Richman, 'On the Self-Reference of a Meaning Theory', *Philosophical Studies* 4 (1953), 69–72, and Paul F. Schmidt, 'Self-Referential Justification', *Philosophical Studies* 8 (1957), 49–54.

[12] I hasten to say, so as not to be misunderstood, that I am not principally concerned in this paper to recommend the formalists's thesis regarding the conventional nature of logical rules. However, viewing such rules in this way serves to highlight the contrast between them and the non-arbitrary and compelling criterion proposed here.

[13] See the author's 'The Idea of a Metalogic of Reference', *Methodology and Science: Interdisciplinary Journal for the Empirical Study of the Foundations of Science and their Methodology* 9 (1976), 85–92. Cf. also 'Phenomenology of the Implicit', *Dialectica* 29 (1975), 174–188, in which, from a phenomenological point of view, referential consistency permits both the identification of "projections" (see below in the text) and their elimination by means of a method of "de-projection", in a logically compelling manner. (For Polish readers, see 'Fenomenologia Tego, Co Implikowane', *Roczniki Filozoficzne* 22 (1974), 73–89.)

[14] A book is now in preparation, supported in part by the Max-Planck-Gesellschaft.

[15] For individual analyses which make use of referential consistency as a criterion of meaning, cf. the author's 'A Metatheoretical Basis for Interpretations of Problem Solving Behavior', *Methodology and Science* 11 (1978), 59–85, specifically §§10, 12;

'Towards a Unified Concept of Reality', *ETC.: A Review of General Semantics* **32** (1975), 43–49; 'Self-Reference, Phenomenology, and Philosophy of Science', *Methodology and Science* **13** (1980), 143–167.

A group of analyses in terms of referential consistency is detailed in *A Relativistic Theory of Phenomenological Constitution: A Self-Referential, Transcendental Approach to Conceptual Pathology* (English and French, 2 vols., Université de Paris 1970 – *Diss. Abs. Internatl.*, No. 79–05.)

[16] The convention is followed whereby False = 0, and the designated truth-value is $n - 1$; the value n is reserved for a purpose described later.

For generality, $p_i s$ with *variable* truth-value may be included: e.g., $p_i s$ for which value assignments are a function of time, as may be the case for future contingent statements, "So-and-so is alive", etc.

[17] It follows from this formulation that a person can refer identifyingly to only one object, of a set of possible objects of reference, at a time. The object referred to may be single or it may be compound, as when reference is made to a set having more than one member, or to a set of sets of objects, etc.

From the perspective presented here, when reference to an object o_i is uniquely determined, o_i is unambiguously identified in the sense of (1) in the text. The *identity* of o_i will essentially be a function of o_i's identifiability – hence, ultimately of frameworks relative to which reference to o_i can obtain.

A good deal must be omitted in this brief treatment: The possibility of re-identification would, for example, as Strawson has pointed out, need also to be assured.

[18] This recalls Quine's dictum, "no entity without identity." (Cf. Leonard Linsky, *Referring* (New York: Humanities Press 1967), p. 27.

[19] On the nature of '$\supset$' in such expressions as '$p_i \supset R\alpha o_i \sigma$', see Note 23 below.

[20] I.e., reference is made by α at σ to the (compound) object of reference $\{o_i, Mp_i\}$.

[21] The expression 'metalogical self-referential inconsistency' need not be restricted to the case in which reference obtains to $\{o_i, M_{p_i}\}$ at a single space-time σ. If $R\alpha o_i \sigma$, and $R\alpha M_{p_i} \sigma'$, σ' is later than σ, and $\bar{p}_i$, then we have the case where α realizes in retrospect that a p_i endorsed by him is projective, i.e., that in endorsing p_i at σ he was metalogically self-referentially inconsistent. Analogously, we may have the case where $R\alpha o_i \sigma$, $R\beta M_{p_i} \sigma'$, and σ' is later than σ: i.e., one man's commitments can be the basis of another man's metalogical analysis.

It is sometimes important to make a similar distinction in connection with pragmatic self-referential inconsistencies. Statements are sometimes and even frequently made by some individuals who are not aware at the time, and may never become aware, of the pragmatic self-referential inconsistencies they involve.

[22] For more detailed illustrations of projective forms of reference, see Note 15.

[23] The reader may be interested in contrasting the variety of entailment in question in this paper with "virtual implication" described by Hintikka, in which '$p \supset q$' is "self-sustaining": See J. Hintikka, *Knowledge and Belief: An Introduction to the Logic of the Two Notions* (Ithaca, N.Y.: Cornell University Press 1962), pp. 32, 57, and passim.

[24] $\mathscr{P}'$ includes $\mathscr{P}$ as a subset; $\mathscr{P}'$ contains in addition to $p_i s$ which fall in the significant range, $p_i s$ which have the value μ.

[25] A three-valued logic, in which the third value is 'meaninglessness' or 'undefined', is

used by Bochvar to stand for the value of paradox-generating propositions. Although his three-valued system is without a theory of types, it is nevertheless consistent. See D. A. Bochvar, 'Ob odnom tréhznačnom isčislénii i égo priménénii k analizu paradoksov klassičéskogo raširénnogo funkcional'nogo iscisléniá' [On a three-valued logical calculus and its application to the analysis of contradictions], *Matématičéskij sbornik* **4** (1939), 287–308; and D. A. Bochvar, 'K voprosu o néprotivoréčivosti odnogo tréhznačnogo isčisléniá' [On the consistency of a three-valued calculus], *Matématičéskij sbornik* **12** (1943), 353–369; as well as Alonzo Church, 'Review of D. A. Bochvar's "On a three-valued logical calculus and its application to the analysis of contradictions"', *Journal of Symbolic Logic* **4** (1939), 98–99; with a correction in *Journal of Symbolic Logic* **5** (1940), 119.

Patrick Suppes makes use of Bochvar's three-valued system (without, however, crediting Bochvar for his truth-matrices) in connection with a formal representation of operationally meaningless statements. Cf. P. Suppes, 'Measurement, Empirical Meaningfulness, and Three-valued Logic', in P. Suppes, *Studies in the Methodology and Foundations of Science: Selected Papers from 1951 to 1969* (Dordrecht-Holland: D. Reidel 1969), pp. 65–79. (Reprinted from C. W. Churchman and P. Ratoosh (eds.), *Measurement: Definitions and Theories* (New York: John Wiley 1959), pp. 129–143.)

Several other authors have proposed three-valued systems in which the third value is 'meaninglessness'. For example: Sören Halldén, *The Logic of Nonsense* (Uppsala: Universitets Årsskrift II, 9 (1949)); Moh Shaw-kwei, 'Logical Paradoxes for Many-valued Systems', *Journal of Symbolic Logic* **19** (1954), 37–40; Lennart Åquist, 'Reflections on the Logic of Nonsense', *Theoria* **28** (1962), Part I, 138–157. For various reasons, however, special properties of these proposed systems make them unsuitable in the present context.

It might be mentioned that some authors have felt that the matrix for negation given in the text precludes a satisfactory interpretation of three-valued logic. That A and $-A$ have the same value when A has the value 'meaninglessness' seems to them problematic. Andrzej Mostowski, for example, has remarked in this connection that he does not have "any hope that it will ever be possible to find a reasonable interpretation of the three-valued logic of Łukasiewicz [which has the same matrix for negation as in Bochvar's system] in terms of ordinary language." A. Mostowski, 'Review of Helen Rasiowa's "A dziedziny logiki matematycznej. II. Logiki wielwartościowe Łukasiewicza"'. [From the domain of mathematical logic. II. The many-valued logic of Łukasiewicz], which appeared in *Journal of Symbolic Logic* **15** (1950), 223. Rasiowa's original paper appeared in *Matematyka* **3** (1950), 4–11.

It is, of course, my belief that Mostowski's pessimism was ill-founded.

[26] For an indication of the rationale behind the condition requiring that re-identification be possible in SI, see the text below, where this Note-number is repeated.

[27] Nothing need be said in any detail about what one's "intentions" may have been in using p_i in this way, since referring to what one had in mind but sees was not realized in actual expression, is in practice to orient oneself with respect to p_i-in-SI in the manner already described.

[28] Cf. Notes 13 and 15.

Suggestions for Further Study

There are a great many works which I would like to mention here. But rather than deluge the reader with a lengthy bibliography, I have decided in favor of mentioning only a small number of contemporary works, in the belief that sometimes fewer choices are more readily acted upon. Classical works in philosophy, by Plato, Hume, Kant, etc., are essential supplements to this short reading list, which quickly branches to many other books, if you make use of the bibliographies contained in the contemporary sources.

Fitch, Frederic Brenton, Symbolic Logic (New York: Ronald Press 1952), in particular, "Self-Reference in Philosophy" (reprinted from Mind), pp. 217-225. This and the following paper by Fitch are two classical contributions to the metatheory of self-referential argumentation.

Fitch, Frederic Brenton, "Universal Metalanguages for Philosophy", Review of Metaphysics, Vol. 17, 1963-64, pp. 396-402.

Hintikka, Jaakko, Knowledge and Belief (Ithaca: Cornell University Press 1962). An interesting account of a group of epistemological questions by an incisive logician.

Hofstadter, Douglas, Gödel, Escher, Bach: An Eternal Golden Braid (New York: Vintage Books 1979). A popular and delightful recreational book which focuses on self-reference.

Johnstone, Henry W., Jr., Philosophy and Argument (University Park, Pa.: Pennsylvania State University Press 1959). Perhaps more than any other philosopher, Johnstone has attempted to work out a theory of philosophical inquiry which makes essential use of self-referential argumentation.

Johnstone, Henry W., Jr., Validity and Rhetoric in Philosophical Argument: An Outlook in Transition (University Park, Pa.: The Dialogue Press 1978).