PARADOX, PREDICTION AND SELF-VITIATING HYPOTHESES

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Alfred Jackson Stenner

1968

This is to certify that the

thesis entitled

PARADOX, PREDICTION AND

SELF-VITIATING HYPOTHESES

presented by

Alfred Jackson Stenner

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Philosophy

William J. Callaghan
Major professor

Date May 15, 1963

O-169

ABSTRACT

PARADOX, PREDICTION AND SELF-VITIATING HYPOTHESES

by Alfred Jackson Stenner

Predictions in the social sciences often have a puzzling effect: the utterance of the explanandum of the prediction brings about a state of affairs which is described by the contradictory or a contrary of the explanandum. These phenomena have been called by a number of names: e.g. "self-stultifying belief", "suicidal belief", "self-destroying hypothesis", "Oedipus effect", etc. The term employed in this essay is "self-vitiating hypothesis".

The occurrence of self-vitiating hypotheses has led some writers to claim that important methodological differences obtain between the social and non-social sciences. It has also been urged that self-vitiating hypotheses make an exact science of history logically impossible. This essay examines several of these related claims and attempts to show that they are not cogent.

In order to deal with the various claims concerning self-vitiating hypotheses, a formal explication of the term is undertaken. This is followed by an analysis of the related predicates "self-vitiatable hypothesis" and "self-fulfilling hypothesis". While the occurrence of self-

vitiating hypotheses is, at present, predominantly a social science phenomenon, such hypotheses may also be problematic for the non-social sciences.

The solution of the problems encountered in dealing with self-vitiating hypotheses requires a theory which will enable us to predict which of our utterances are, and which are not self-vitiating. At present no such theory is available and the difficulties which will attend the development of such a theory are formidable. But the difficulties are empirical and not logical.

If and when a theory of self-vitiation is developed, other puzzles may arise. The theory of self-vitiation will not necessarily be immune to the kind of difficulties which prevent the social scientist from accurately predicting the future. Predictions about certain utterances being self-vitiating may themselves turn out to contain self-vitiating explananda. These difficulties are shown to be empirical, and they may be resolved by the development of a more comprehensive theory of self-vitiation.

Predictions of one's own behavior may also contain self-vitiating explananda. While the problems encountered in these instances differ in some respects from problems considered previously, they are not beyond the possibility of resolution.

The essay concludes with a brief examination of some of the relations which may obtain between a theory

of values and a theory of knowledge due to the phenomena of self-vitiating hypotheses.

PARADOX, PREDICTION AND SELF-VITIATING HYPOTHESES

by

Alfred Jackson Stenner

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Philosophy

1963

I believe in order that I may know. -- St. Anselm

If you don't make any assumptions, you can't draw any conclusions. -- Henry Leonard

If we are ready to tolerate everything as understood, there is nothing left to explain; while if we sourly refuse to take anything, even tentatively, as clear, no explanation can be given. -- Nelson Goodman

PREFACE

The following essay concerns the phenomenon of self-vitiation, a subject of considerable importance for the philosophy of science. Yet it is a kind of phenomenon which has received but scanty attention in the literature. The fact that self-vitiation is often overlooked in discussion of science may be due in part to the fact that we have so little understanding of the phenomenon in question. It is hoped that this essay will contribute to an increased awareness of the problems which are posed by self-vitiating utterances, and also to an increase in time and other resources which will be required in attempting to resolve some of the problems in this area of inquiry.

I am grateful to the members of my guidance committee: William J. Callaghan, Ph.D., Paul M. Hurrell, Ph.D., Richard S. Rudner, Ph.D., and John F. A. Taylor, Ph.D. for their helpful comments and encouragement.

Those who are familiar with the recent philosophical literature will recognize my indebtedness to Professors Nelson Goodman, Carl Hempel, Henry Leonard, and Paul Oppenheim.

Above all, I wish to thank Professor Richard S. Rudner who has guided this work from its inception to its completion. My conversations with him about the material in this essay have all proved singularly rewarding. His comments and suggestions have brought a degree of organization and clarity to the following material which would not have been otherwise achieved.

Finally, I wish to express my gratitude to my family, my wife, Jean, and our three sons, Jack, Jim and Rob, without whose encouragement, sympathy and tolerance through a trying period, this manuscript would never have been completed.

TABLE OF CONTENTS

CHAPTER I

A PRESYSTEMATIC DISCUSSION OF SELF-VITIATION

Introduction

One of the most important function, if not the most important function of scientific knowledge is that it enables us to predict, with a high degree of success, what the morrow will bring. Sometimes, however, the very utterance of the prediction seems to bring about an event which is described by the contradictory of the utterance. In this chapter we shall adopt a terminology which will enable us to describe, with modest precision, the phenomena in question, and then give some examples of the phenomena which will give some indication of the range of application of the predicate 'self-vitiating'.

The Nature of the Problem

Throughout the recent philosophic literature, a particular argument -- or what could be considered as a set of related arguments -- has received considerable attention. The claim has been made that methodology and content of the social sciences differs in certain specifiable ways from the methodology and content of the physical sciences. And it is further claimed that these differences are of such a nature that for purposes of methodological analysis it makes little or no sense to speak, in Quine's

phrase, of "The whole of science". Among the concerns with which the present essay will be concerned is whether the proponents of this view have made good their claim.

Not all of the arguments which have been advanced in favor of this thesis will, of course, be dealt with in what follows. We will not, for example, be specifically concerned with the arguments which have frequently been made from verstehen; nor will we be concerned with arguments concerning the uniqueness of, or peculiarly teleological character of, social science phenomena. These arguments have been effectively answered elsewhere[1] and need not detain us, except to note that it is this author's conviction that each of these arguments suffers from a fatal flaw.

The argument to which we shall direct our attention, has appeared less frequently in the literature. It concerns a class of phenomena which has been called by a number of names to characterize a peculiar feature shared by all

---

[1] For a criticism of verstehen as a methodological principle, see Theodore Abel, "The Operation Called Verstehen", American Journal of Sociology, 54:211-218, 1948-1949, reprinted in Edward W. Madden, The Structure of Scientific Thought (Boston: Houghton Mifflin Co., c.d. 1960) pp. 158-166. For a discussion of the place of law-like statements in historical explanation, see Ernest Nagel, The Structure of Science, (New York: Harcourt, Brace, and World, Inc., c.d. 1961) pp. 547-606. An extensive bibliography for treatments of these issues is found in Madden, op. cit., pp. 372-376.

members of this class. These have been called "self-destroying prophecies",[2] "suicidal predictions",[3] "self-destroying beliefs",[4] while Richfield and Copi refer to them as "predictions...which are vitiated".[5] We shall employ the predicate 'self-vitiating hypothesis' to characterize the phenomena in question. One of the important facets of the explicandum we are here considering is its suprising recalcitrance to very clear explication. Accordingly, a major effort in the discussion to follow will be given over to an attempt to furnish a fairly rigorous and precise explicatum for this term. For the present, however, we shall begin with an informal characterization of the phenomena to which we intend to apply the term and then give a number of illustrations which indicate the scope of the

---

[2]Alan Gewirth, "Can Men Change The Laws Of Social Science", Philosophy of Science, 21:229-241, July, 1954.

[3]Ernest Nagel, op. cit., pp. 468-473.

[4]R. K. Merton, Social Theory and Social Structure, (Glencoe, Ill.:The Free Press, 1957, Revized and enlarged edition.) pp. 121-130.

[5]Jerome Richfield and Irving M. Copi, "Deciding and Predicting", Philosophy of Science, 28:47-52, January, 1961.

term's application.

Adoption of a Uniform Terminology

In order to keep our presystematic discussion of self-vitiation as free of ambiguity and vagueness as possible we shall find it convenient to adopt a uniform terminology.

By hypothesis we shall mean any statement or utterance of a statement.

We shall use the term 'statement' neutrally to avoid any unnecessary commitment to the existence of propositions, meanings and like entities.[6] Thus construed, the term 'statement' may be understood as "a declarative sentence-type insofar as it is used to indicate a certain proposition"[7] or as an individual composed of a number of sentences "catalogued under a single label",[8] i.e., an individual composed of all the replicas of a particular sentence or the sentential part of an individual sign-event.

---

[6]The reasons for this decision are given below.

[7]Henry Leonard, Principles of Right Reason (New York: Henry Holt and Co., c.d. 1957), p. 615.

[8]Nelson Goodman, The Structure of Appearance (Cambridge, Mass., Harvard University Press, 1951), p. 290.

The following analysis shall be of such a nature as to be serviceable in a system which interprets 'statement' in any of the ways described above. At least, such is our hope. It should be clear, however, that such a hope will be fulfilled only if the cogency of our construal of the term 'statement' is not dependent on (but is compatible with) the "existence" of propositions or meanings. Hence, although we claim to use the term 'statement' neutrally, this claim should not be taken to mean that we are using the term ambiguously.

In particular, we shall use the term 'statement' so that it may be interpreted in the nominalistic sense indicated by Nelson Goodman,[9] i.e., any number of sentences which may be construed as catalogued under a single label will be called replicas of one another; and all the replicas so included in a catalogue under a single label may be construed as a spatially and temporally discrete individual which we call 'a statement'. We are contending that such an explication of the term'statement' will also be adequate to a so-called platonistic system which recognizes propositions and meanings as values of its

---

[9] loc. cit.

variables. It is therefore only in the sense just described that our usage of 'statement' is to be understood as neutral.

It will be necessary to recognize two different usages of the term 'hypothesis'. According to one usage we shall understand simply an individual utterance or a number of utterances of a particular statement. When we wish to symbolize a particular hypothesis in this sense we shall use the lower-case letter '_h_'. When we wish to speak of an _hypothesis_ in the sense of a _statement_ (i.e. as an individual composed of _all_ the replicas of a given sentence) we shall use the upper-case letter '_H_' . When the distinction is not required for the purposes of the analysis, or when the context makes it clear which of the two usages is being employed we shall speak simply of an hypothesis. When, in the discursive portions of the text, the distinction is required, we shall speak, on the one hand, of _utterances_ of an hypothesis or _utterances_ of a statement, and, on the other hand, of statements of an hypothesis.

We shall use the term '_utterance_' or '_statement-utterance_' to refer to particular sign-events in which a replica of a given sentence is produced. We shall use the term '_sentence-token_' or '_tokens of a particular_

statement' or 'inscription' to refer to the physical concrete objects so produced.

The term 'prediction' calls for extended comment. We shall use the term to refer to those linguistic entities which conform to the requirements for constituting a prediction laid down by Carl G. Hempel and Paul Oppenheim.[10] In particular, only those entities will be considered as predictions which fulfill the following requirements:

1. The explanandum must be a logical consequence of the explanans.
2. The explanans must contain general laws, and these must actually be required for the derivation of the explanandum.
3. The explanans must have empirical content.
4. The sentences constituting the explanans must be true.

The above requirements do not avoid all the difficulties involved in determining what does and does not constitute a scientific explanation or prediction. But the Hempel-Oppenheim paradigm is, at present, the most definitive statement available. Consequently, we shall perhaps be justified in adhering to these requirements

---

[10]Carl G. Hempel and Paul Oppenheim, "Studies in the Logic of Scientific Explanation", Philosophy of Science, 15:135-175, April, 1948.

until a more definitive statement is forthcoming.[11] What we are explicitly affirming is the logical symmetry of explanation and prediction. The characteristic which distinguishes prediction from explanation is a temporal one. A prediction contains an explanandum which is actually deduced and uttered prior to the event described by the explanandum. An explanation contains an explanandum which is deduced and uttered after the event described by the explanandum.

A large class of utterances which shall concern us do not conform to the requirements of the Hempel-Oppenheim thesis and yet are of such a nature as to be of concern to the social scientist. These utterances may be shrewd guesses, educated guesses, prophecies, "predictions" (in the vulgar sense) and like utterances. These utterances are of concern to the social scientist since they too may have a self-vitiating effect, presystematically understood. These utterances we shall refer to as 'prophetic-hypotheses' or as '"predictions"'. The use of double quotation marks serves to indicate that the entities in question do not conform to requirements 1-4. Those statements which can

---

[11]For a recent defense of the Hempel-Oppenheim thesis, see Adolf Grunbaum, "Temporally-Asymmetric Principles, Parity between Explanation and Prediction, Mechanism Versus Teleology", Philosophy of Science 29: 146-174, April, 1962.

serve as explananda of scientific predictions we shall call 'fore-dictions'.

Another class of utterances needs to be considered. This kind of utterance, at first glance, appears to be one which should be subsumed under those labeled 'fore-dictions' but which on further analysis turn out not to be deducible from an explanans for one of the following (or other) reasons: they contain mispelled words, or are grammatically incorrect, or are amphibolous or contain simple ambiguities or vague expressions. Hence we shall treat these malformed hypotheses as a sub-class of prophetic-hypotheses. In Chapter III we shall consider ways of dealing with prophetic-hypotheses which we would presystematically understand to be self-vitiating. But the interesting cases which shall be our especial concern are fore-dictions. We shall therefore restrict the predicate 'self-vitiating' to an hypothesis which has the following characteristics:

a) Predictions of which it is an essential component (i.e. the explanandum) are not borne out when the hypothesis is uttered.

b) The hypothesis and other circumstances are such that we are warranted in asserting the counterfactual: "If the hypothesis had not been uttered, then the prediction would have been fulfilled".

The requirements 1-4 listed above together with the informally indicated defining characteristics a) and b) serve at the outset to mark off a reasonably precise area for investigation. It is therefore hoped that the ensuing discussion will be marked by a rigor and precision which has been lacking in some of the discussions in the literature.[12] We may note how some of the vaguenesses and ambiguities which have characterized such terms as 'suicidal belief' and 'self-destroying prophecy' are removed by these requirements. Characteristics a) and b) serve to discriminate between those hypotheses which we want specifically to consider and those which are simply false. While we may intuitively grasp the distinction between the two types of hypotheses we shall have occasion to note, as was remarked above, that the formal characterization of the term 'self-vitiating hypothesis' poses some difficult problems.

Requirements 1-4 enable us to distinguish between "shrewd guessing" or "prophesying" on the one hand, and predicting, on the other. It may be the case that guessing and prophesying also may be self-vitiating in some

---

[12] See Chapter II for a criticism of certain attempts to deal with the problem of self-vitiation.

presystematic sense of the term. If so, then the features which characterize self-vitiating fore-dictions will also be features of certain guesses and prophesies. But while guessing and prophesying about the future may often turn out to be either felicitous or to share other characteristics of fore-dictions, they have no scientific status. The peculiar puzzles created by certain prophetic-hypotheses shall be examined and dealt with in Chapter III. To repeat, the predicate 'self-vitiating' shall be reserved for fore-dictions.

Merton's use of the term 'self-destroying belief' is particularly ambiguous due to the difficulty of determining the precise character of the phenomena in question. One of his examples illustrates this particular difficulty:

> In the dark days of 1862, when McClellan was stalemated and the armies in the west immobilized, Lincoln did not issue a public call for the desparately needed thousands of new troops, explaining, "I would publicly appeal to the country for this new force were it not that I fear a general panic and stampede would follow, so hard is it to have a thing understood as it really is".13

---

[13]Merton, op. cit., p. 128.

This particular passage appears in the section, of Merton's book, dealing with self-destroying beliefs. What is unclear from the illustration and Merton's treatment is exactly what is being destroyed or would be in danger of being destroyed had Lincoln issued his call for troops. Would it have been the Union? The morale of the northern troops? The morale of the citizens in the northern states? Or would it have been Lincoln? His prestige? This particular kind of difficulty is obviated if we make clear from the outset the kind of entities that shall be our concern. The kind of entities which shall concern us are linguistic; and they have the peculiar property or characteristic of being causally related to the event which is described by their negation.

We need to discriminate even further and distinguish statements from utterances of statements or statement-occurences. Likewise we need to distinguish those utterances of statements which enter into a causal relationship with the event described by the negation of that statement and those utterances of the same statement which do not enter the causal relationship. An example may show the need for these distinctions. Suppose that a particular statement S describes e and that as a result of certain utterances of S, ~e comes about. Now as we have defined

'statement' we may note that it is not all the utterances of S which are causally efficacious in bringing about ~e but only some of them. If, for instance, some of the utterances of S occur after ~e we shall probably not wish to say that these later utterances entered the causal sequence which brought about ~e.[14] A more detailed elaboration of these points will be undertaken in Chapter III. At this point we need only remark that the predicate 'self-vitiating' will be applied to utterances of statements. Utterances, it should be remembered, are construed as sign-events. And hence utterances may function as values of variables in a causal explanation.[15]

One further distinction needs to be noted. We distinguish an utterance of S, here construed as a psychosocial event, from the sentential components or parts of the utterances of S. The latter are not events but physical objects which are parts or constituents of events. Following Nelson Goodman, we may draw the distinction by

---

[14]The only cases in which this particular assertion will not hold will be those in which all the utterances of S predede ~e and each and every utterance of S enters the causal relationship.

[15]Our construal of the term 'causal explanation' is delineated in Chapter III.

noting that utterances contain times whereas their sentential parts do not.[16] This is not to suggest that sentence-tokens are not fully concrete entitites. It is to suggest that for the purposes of analysis we attend only to those qualia of tokens that are other than time-qualia. We could have treated sentence-tokens as containing time-qualia and hence as events. And presumably an analysis operating on this assumption would also prove adequate. It may therefore seem that our method of distinguishing between utterances and tokens is somewhat arbitrary. It is arbitrary only in the sense that we need to make some decision concerning what may serve as values of the variables in any adequate theory of self-vitiation. And it is entirely possible that an analysis which allowed tokens (here construed as events) to serve as the value of the variables in a theory of self-vitiation would also be adequate for explaining the phenomena in question.

We shall attempt to avoid making essential application of the predicate 'self-vitiating' to such entities as meanings, propositions, states of affairs, and statements (when these are construed as abstract entities), although

---

[16] op. cit., pp. 283-287.

it appears likely that so nominalistic an orientation may become exceedingly difficult to maintain in attempting to resolve some of the puzzles posed by the phenomena we are examining. To some, our reasons for imposing such rigid restriction on our analysis may be a source of wonder. Why, it may be asked, when meanings play so obvious a role in the phenomena under consideration would one be constrained to avoid reference to meanings when such reference could presumably aid in the analysis? In reply we may say that abnegations in philosophy, no less than in religion and morals, have their own peculiar rewards. The basic notion here is that of simplicity. We wish to keep our universe as uncluttered as possible. But at the same time we wish to be able to say all that is worth saying concerning the phenomena in question. What we shall therefore attempt is to show how we can deal with the designative role of utterances in relationship to the events they purport to describe without introducing 'meaning' as a primitive or undefined term. We do not ignore the question of meaning, but for the sake of simplicity we shall attempt to avoid giving "meanings" -- and other similar entities -- ontological status. In other words, meanings will not serve as values of the variables of lowest type in any system

based on the following analysis. If we can carry this plan through to completion, our analysis will be simpler than one which allows the introduction of a greater number of entities as values of its primitives.

Some Illustrations of Self-Vitiating Hypotheses

The phenomena in question are fairly common occurrences in daily life. And the following illustrations are not meant to exhaust the various types of hypotheses in question but only to suggest the kind of entities with which we are concerned in this essay.

A. An economist predicts, on the basis of reliable information and highly confirmed economic laws, that, due to a shortage of a particular commodity C, companies X, Y, and Z which are presently engaged in producing C will realize a profit P at time t.[17] Subsequent events, however, show that the explanandum of the prediction turned out to be false. Those companies which had expected, on the basis of the prediction, to make a profit, instead sustain a loss. What has happened is the following: when the fore-diction was itself made a piece of public information, other companies, hoping to profit by the production of C began

---

[17] Throughout the remainder of this discussion, the requirements that what are acceptable as predictions are those which conform to points 1-4 on page 7 of this essay will be tacitly understood and will not be reiterated. An important qualification is made in Chapter V, p. 154.

flooding the market with this item. The result is that all the companies engaged in producing C are faced with large inventories and loss of profits. Wherein lies the difficulty with the fore-diction? If we are warranted in asserting the counterfactual, "If the hypothesis (i.e. the explanandum) had not been uttered, then companies X, Y, and Z would have made a profit at time t." we may predicate 'self-vitiating' of this particular hypothesis.

B. The National Safety Council publishes a statement to the effect that the Council predicts at least four hundred Americans will die as the result of automobile accidents during a particular holiday, Hol. This statement is published in newspapers, and carried on spot television and radio announcements. The explanandum of the prediction is not borne out by a subsequent examination of the facts. We may "explain" this sequence of events by noting that the prediction may have had the observed effect through its own allusion to the dangers involved in travelling during Hol.; i.e., many Americans, learning of the fore-diction, had remained at home, and others had exercised more caution than usual during their travels. From the evidence it seems plausible to assume that predictions by the National Safety Council often do have just such an effect. That their fore-dictions do not have this

self-vitiating character more often may be due to the fact that those who make the predictions hope that their campaign, coupled with the utterances of their hypotheses, will actually have such a self-vitiating effect. It may also be the case that those responsible for making the predictions actually do take account of the possible self-vitiating character of such fore-dictions and revise their predictions accordingly.

C. At times this self-vitiating effect may be used with deliberate intention to deceive. A commander of a regiment in time of war predicts, on the basis of information gathered from prisoners, espionage, etc. that the enemy is about to make a move M. Discovering that he is ill-prepared to defend against M, he allows his fore-diction to "leak" back to the enemy camp in the hope that his opponents will abandon M as a plan of operation.

D. An economist predicts that a certain event E will take place relative to conditions of the stock market. He publishes his prediction in one of the leading economic journals; and the statements describing the future occurence of E are thence published in the Wall Street Journal where they are subsequently read by most of the major investors in the stock market. As a result of the prediction, the activity of the investors is altered. E does not occur.

Common sense would again lead us to assert the counter-factual, "If the fore-diction had not become known to the buying public, E would have taken place."[18]

E. In 1948, most of the major public opinion polls predicted that Thomas E. Dewey would become the next President of the United States. That this fore-diction was not borne out is a fact of history. Some of those who had supported Mr. Dewey blamed the pollsters for engaging in "pseudo-science" and basing their predictions on insufficient evidence and poorly confirmed laws. A few voices, however, suggested that one possible cause of the outcome of the election had been the effect of the utterances of the fore-diction itself on Republican voters who simply stayed away from the polls, convinced as they were of a Dewey victory. The minority voices were saying in effect that the fore-diction was self-vitiating; i.e. they were asserting the counterfactual, "If the fore-diction had not become public information, Mr. Dewey would have won the election." If we raise the question as to why the predictions of presidential elections in 1952, 1956, and 1960 did notshare this self-vitiating characteristic,

---

[18]This example is a paraphrase of one suggested by Popper in Poverty of Historicism.

common sense has a ready answer: "The general public has become aware of the self-vitiating character of such fore-dictions and has acted accordingly."[19] We shall later have occasion to note that the foregoing answer of common-sense leads to some puzzling consequences.[20]

---

[19] An interesting feature of the 1960 presidential election, however, concerned just this self-vitiating feature. N.B.C. and C.B.S., which televized the election returns, made use of gigantic computers which determined the odds of a Kennedy or a Nixon victory. A few minutes after the first returns began to be fed to the computers, the odds in favor of a Kennedy landslide were seven hundred to one. These odds remained for almost an hour after the polls had closed in the East. In the West, however, the polls remained open for three hours after they had closed in the eastern quarter of the nation. When the polls had finally closed in the West and the votes had been tabulated, Kennedy's odds had been so reduced that the difference in popular votes cast for him and Mr. Nixon was but a few thousand. That the computer's "predictions" had such a self-vitiating effect seems plausible, even in the absence of empirical evidence.

[20] For one thing it tends to vitiate the self-vitiating character of certain fore-dictions. For another, it seems to be based on some general law of human behaviour which may be roughly phrased: "Whenever the general public becomes aware of the self-vitiating character of a prediction, it will act in such a way as to vitiate the self-vitiating character of the prediction; or it will act in such a way as to bear out the initial prediction." Such general law-statements may raise puzzles in their own right. On the basis of such a law, we can make certain predictions. But then we will wish to know whether utterances of this law and hypotheses deducible from it are self-vitiating. See Chapter VI for an extended discussion of this problem.

F. Karl R. Popper has argued that it is precisely this self-vitiating character of our predictions about the future course of human history which makes such predictions "logically impossible". At this point, we may simply take note of his argument which in substance is as follows: Knowledge is for doing. And insofar as what we do is dependent upon our knowledge about the world, just so far is it necessary for us to predict what knowledge we shall possess in the future if we are to successfully predict our future behavior. But it is not logically possible to predict our future knowledge; since to predict that at some future time, t, we would have a specific piece of knowledge K, would be to vitiate the prediction. This would be due to the fact that we would need to possess K prior to t in order to make the prediction. Hence all predictions about action which is dependent upon knowledge of what our future knowledge will be is in effect self-vitiating.[21]

While Popper's arguments have an initial plausibility, we hope to show that they will not bear examination; and that much of the plausibility is due to an ambiguity in the term 'knowledge'.

---

[21]Karl R. Popper, Logic of Scientific Discovery, (London:Hutchinson, 1959).

G. Another, and perhaps related example is offered by Richfield and Copi.[22] "Suppose I agree to make a decision between alternative courses of action A and B next Wednesday, and that I have promised to delay my action until that day." Suppose further that on the basis of knowledge about myself, my previous decisions, certain tendencies to act in a certain way under certain speciable circumstances, I predict that I will decide in favor of A on next Wednesday. One might argue that to make such a prediction is in effect to make the decision now and hence to vitiate the fore-diction.[23]

Professor Henle, for example, has argued[24] that "any prediction I make about myself is vitiated by the fact that I make it...While I may predict the weather or coming political events or your conduct, I do not predict at what time I shall have lunch or go to bed. These are questions which I simply decide." Here the notion of predicting

---

[22]Jerome Richfield and Irving M. Copi, "Deciding and Predicting" Philosophy of Science, 27:47-51, January, 1960.

[23]Although the example is due to Richfield and Copi, they are concerned to deny that the prediction is self-vitiating.

[24]Paul Henle, "Do We Discover Our Use of Words?" Journal of Philosophy, 54:750, November 7, 1957.

events in daily life requires considerable sharpening before we can examine the logical issues involved.[25]

H. Another example:[26] Nation N is at war with nation M. N has all of the military advantages at its disposal; it possesses money, manpower, and arms together with the means for producing more of everything than does M. M is lacking in financial resources; its manpower is depleted; it lacks food and arms in sufficient quantities to wage a successful war against N. The odds for a victory by N are so overwhelming that social scientists begin making predictions to the effect that N will win the war. The causal consequences of the utterances of the social scientists are the ensuing lethargy and lack of interest on the part of the populace of N in regard to the war effort and renewed determination on the part of the populace of M to win the war. The final result: M wins the war. In such circumstances we believe ourselves to be warranted in asserting the counterfactual: "If the

---

[25]Thus the notion of prediction here seems not to conform to the requirements indicated above although it may do so in specific instances.

[26]This example is a paraphrase of one given by Alan Gewirth, op. cit.

fore-diction, 'N will win the war.' had not been uttered, N would have won the war."

John Venn was one who held that this class of phenomena enables us to distinguish between the social and the physical sciences. He writes:

> Any person can see that to draw inferences about a thing, and then to introduce a disturbance which was not contemplated when the inference was drawn is to invalidate the conclusion we have obtained. But when the inference is about the conduct of human beings it is often forgotten that in the inference itself, if published, we may have produced an unsuspected source of disturbance. In other words, if the results of our investigations be given in the form of statements as to what people are doing and what they will do, the moment these statements come before their notice the agents will be subject to a new motive, which will produce a disturbance in the conduct which had been inferred...
>
> The publication of the Nautical Almanac is not supposed to have the slightest effect upon the path of the planets, whereas the publication of any prediction about the conduct of human beings (unless it were kept out of their sight, or expressed in unintelligible language) almost certainly would have some effect. The existence of this distinction renders all such physical illustrations entirely inapplicable when we thus attempt to explain the way in which it is supposed that human conduct can be studied and foretold.[27]

---

[27]John Venn, Principles of Empirical or Inductive Logic (London, New York: MacMillan and Company., 1889), pp. 575-576.

Venn's general position seems to be one which is popularly subscribed to, although it is one which, as we hope to show, will not prove to be completely satisfactory; nor are all of his arguments rendered plausible when subjected to careful analysis.[28]

The Relevance of the Issue

The issues raised by the foregoing discussion may be discussed under two broad headings which, loosely speaking, may be labeled "practical" and "theoretical". We shall simply take note of the former and throughout the course of our discussion it will be tacitly assumed that the latter is the point of our inquiry. This is not to suggest that the practical aspect of the issue is of less importance in the total scheme of human affairs than the theoretical. For it is entirely possible that a self-vitiating hypothesis could have disastrous historical consequences. The point is rather that we would prefer to leave empirical questions concerning any particular self-vitiating hypothesis or set of hypotheses as issues with which the sociologist, psychologist and historian are better equipped to deal. We may note that the kind

---

[28] In Chapter II we shall have occasion to examine Venn's position in greater detail.

of examples offered by Merton[29] and Gewirth[30] are sufficient to dispel any belief that self-vitiating hypotheses are of little moment in regard to the concerns of everyday life. The practical politician and statesman appear to be well aware that predictions or fore-dictions on which public domestic policy and international diplomacy are based would have disastrous effects if they became known prior to the time of the occurence of the events predicted. And much of the secrecy surrounding the activities of the State Department in Washington and its various counterparts throughout the world can be understood or "explained" not simply as attempts to guard military secrets and scientific information; but also as attempts to prevent other nations from discovering what we "predict" they will do in the future. Having noted the practical relevance of the phenomena of self-vitiation, we may leave the task of explaining instances of this type of phenomena to the social scientist.

Our principal concern will be with the relevance of this class of phenomena for epistemological or "theoretical" considerations. We need first of all to note that

---

[29] See p. 11 of this essay.

[30] See pp. 22-23 of this essay.

a part of the program of empiricism from John Stuart Mill, through the Vienna Circle, and up to the present time has been the attempt to show that our knowledge about the world is not simply comprised of isolated facts and theories, but that there is, in Quine's phrase, such a thing as "the whole of science". We need not engage in the debate as to whether this scientific knowledge constitutes the "whole of knowledge" but rather whether this scientific knowledge consists (even ideally) of a unified and integrated whole. If it should turn out that self-vitiating hypotheses constitute a class of phenomena with which an empiricist understanding of science is incapable of dealing, and moreover that such hypotheses are indigenous to the social sciences, then those who hold that there are methodological differences between the social sciences and the physical sciences will have additional ammunition for their guns. If, however, we can show, as we hope to, that self-vitiating hypotheses offer no insuperable obstacle to an empiricist program, then we shall have made some small contribution to the task of showing that the claim for the unity of knowledge about the world is not discredited on these grounds. Specifically we hope to show that the claim that self-vitiating hypotheses are social phenomena which enable us to mark a distinction between the social and the

physical sciences is untenable. In a more positive vein we shall attempt to clarify some of the issues which will need to be considered in the development of a theory of self-vitiation.

CHAPTER II

AN EXAMINATION OF SOME OF THE ARGUMENTS IN THE LITERATURE

Introduction

Although the phenomena of self-vitiation have seldom received detailed analysis from an epistemological or logical point of view, the argument that this kind of phenomena is indigenous to the social sciences has not been without its supporters. In this chapter we shall examine some of the arguments which have been advanced in favor of this thesis and attempt to evaluate the cogency of these arguments.

The Arguments of John Venn

The issue as to whether or not self-vitiating hypotheses give us adequate grounds for marking a distinction between the social and the physical sciences is at least as old as the late nineteenth century. John Stuart Mill believed and argued that, while he had little hope for the scientist's ability to predict with any degree of accuracy individual behavior, he did believe that with the advancement of the social sciences it would be possible to predict human behavior on a large scale or in the mass, i.e., general trends in social movements.

Mill had no doubts or reservations regarding the possibility of making sociology as exact a science as astronomy.[1]

John Venn took issue with Mill precisely on the point of self-vitiating hypotheses:[2]

> If the results of our investigations be given in the form of statements as to what people are doing and what they will do, the moment these statements come before their notice the agents will be subject to a new motive, which will produce a disturbance in the conduct which had been inferred. We may make what statements and criticism we please about the past conduct of men, but directly we commit ourselves to any statements about the future, or, in other words, begin to make predictions, we lay ourselves open to that difficulty just mentioned. The predictions can be made seems to be held by many of those who have adopted the application of logic now under consideration. They do not, of course, claim to be able to foretell the particular actions of individuals, but they assert that it is quite possible that we may some day be able to foretell general tendencies, and the results of the conduct of large masses of men.

Between Mill and Venn we see the battle line clearly drawn between those who claim that prediction of human behavior is often self-vitiating and therefore impossible and those who claim that prediction is

---

[1]John Stuart Mill, A System of Logic (New York: Harper and Bros., Publishers, 1874), pp. 550 ff.

[2]John Venn, op. cit., pp. 575-576.

always possible in spite of ostensibly self-vitiating characteristics. Lest the discussion be obscured by unnecessary ommissions we hasten to add that although Mill and Venn do not explicitly say so they are both discussing successful prediction. Neither of them would wish to deny that from a particular set of law-like statements and well evidenced statements describing antecedent conditions it would be possible to deduce certain conclusions according to valid rules of inference. The point at issue is whether utterances of such deductions are of such a nature that in principle such predictions cannot be successful.

In regard to Venn's arguments several points need to be noted. First, it is not clear in what sense an inference's conclusion may be invalidated by subsequent events. The purport of his objection seems clear enough. But the ambiguity in the term 'inference' makes suspect his precise point. Suppose we have a set of law-like statements L and a set of statements C describing certain antecedent conditions and that the conjunction of L and C logically imply E. We might say then that from L and C taken in conjunction we can infer E. In what sense can this inference be invalidated? Surely not in the logical sense of the term 'to infer'. For no matter

what we add to the premises of the argument, E will still be deducible from the conjunction of L and C. It will be recalled that Venn said "Any person can see that to draw inferences about a thing, and then to introduce a disturbance which was not contemplated when the inference was drawn is to invalidate the conclusion we have obtained."[3] In what sense is Venn holding that the conclusion is invalidated? Surely it is not qua deductive inference. Venn as an able logician would hardly have been guilty of such an elementary blunder.

Another initially plausible construal of his claim might be that it is directed to inductive inference in which, by adding an additional premise to the set, we invalidate the conclusion. Thus, for example, a set of premises 'Fa', 'Fb', ... 'Fn' is confirmatory for the hypothesis '(x)Fx' while one statement such as '~Fy' will tend to disconfirm the hypothesis, ceteris paribus. But consider the following inductive arguments:

---

[3] Loc. cit.

A	B
Fa . Fb ... Fn	Fa . Fb ... Fn . ~Fy
∴ (x) Fx	∴ (x) Fx

Now if we are interpreting Venn correctly he seems to be suggesting that argument A occurs at a time t and that an event ~e occurs at a time later than t, a time which we may call t_0. We assume that the event ~e is described by the statement '~Fy'. Argument B presumably occurs at a time later than t_0 say t_1. Argument B we assume to be invalid by any standards of inductive validity. But in what sense does this render the conclusion of argument A invalid? Yet if we understand Venn's argument he seems to be suggesting that since the premise of B, i.e. '~Fy', is presumably true (since it describes ~e), therefore the conclusion of argument A which occured prior to ~e is invalidated. But this certainly appears to strain the sense of 'validity' to a point where the term will fail to apply to anything.

Perhaps what Venn had in mind when using the term 'invalidate' was simply that the conclusion of argument A is false. In this case his argument loses much of its force. For the scientific status of the conclusion to argument A will depend on a number of factors; e.g. whether or not we are willing to accept as true the

statement '~Fy' and whether or not we are willing to give up the generalization '(x)Fx' after considering the cost involved if we were to abandon it.

A second difficulty arises in regard to Venn's use of the term 'statement'. He says, "The moment these statements (i.e. self-vitiating hypotheses) come before their notice, the agents will be subject to a new notive." The question as to what constitutes a statement and whether or not it is the statement or the occurrence of a statement which is involved is a matter of some importance. A statement-occurrence can be considered as a psycho-social event having certain causal consequences. If, however, what is intended is the assertion that "meanings" are causally efficacious, then certain ontological questions would need to be raised and answered before the problem could even be clearly articulated. What we shall suggest is the the important issues in regard to the phenomena in question can be answered without raising ontological questions of this kind. For the present we may simply observe that Venn is not clear, in his account of the matter, what he intends us to understand by the term 'statement'. And until this ambiguity is resolved it is difficult to assess the weight of his argument.

Perturbance Effects in the Physical and Social Sciences

Up to this point we have been overlooking the major point which Venn wished to make: i.e. that whereas such publications as the Nautical Almanac are not supposed to affect the orbits of the planets, publications about the behavior of human beings certainly do have some effect. Phrased somewhat differently what is being claimed is that fore-dictions in the social sciences have self-vitiating instances whereas fore-dictions in the physical sciences do not. We must now attempt to determine whether or not this claim can be cogently upheld.

Using Venn's example we may ask whether the publication of the Nautical Almanac could have self-vitiating instances. More specifically we may inquire whether any predictions or fore-dictions contained within the Almanac can have replicas such that they are constituents of a sign-event which is causally efficacious in bringing about an event whose occurence disconfirms the prediction or fore-diction.

Prior to the space age we more than likely would have been content to take Venn's arguments at face value. Today, however, with the unleashing of the power of the atom and the possibility of space flight, there does not seem to be evidence to support the claim that the orbit

of Mars could not be altered. Consider, for example, a particular statement S which describes the elliptical path of Mars around the Sun. Suppose that replicas of S are produced by exploding cobalt bombs propelled by rockets into the vicintiy of the planet Mars. These bombs explode in such a fashion as to spell out replicas of S and have the causal consequence of changing the orbit of Mars. It follows that the fore-diction concerning the orbit of Mars is self-vitiating. Now a number of objections may occur to the reader regarding the unconventional tokens employed and the failure to take cognizance of the difference between the use of language for communication and its use for moving the orbits of planets. Many of these objections we hope to consider in later chapters. At the present time we may simply note that the example conforms well with our presystematic discussion of what shall be called a self-vitiating hypothesis. That is, S is deducible from an explanans which conforms to requirements 1-4 listed on page 7. Also, predictions of which it is an essential component, i.e. the explanandum are not borne out when uttered (although, to be sure, this is a somewhat unusual mode of utterance). And finally the hypothesis and other circumstances are such that we are warranted in asserting

the counterfactual: "If S had not been thus uttered then the prediction would have been fulfilled." Hence S fulfills all the requirements for the label 'self-vitiating'.

Now the point of our example needs to be kept clearly in mind. We are not claiming that there is no distinction between certain kinds of self-vitiating hypotheses. At this stage we are only concerned to show that Venn's example, and the claim based upon it, lack the cogency usually attributed to it.

Ernest Nagel gives an example of self-vitiation which is perhaps intuitively closer to the kind of phenomena to which we would ordinarily be willing to apply the predicate. His example also, however, is taken from the physical sciences:

> The pointing and firing of an anti-aircraft gun can be effected by means of a purely physical mechanism. Such a mechanism includes, we may assume, a radar for locating the target, an automatic computer for calculating the direction in which the gun should be pointed to hit the target as reported by the radar, an adjusting device for pointing and firing the gun, and some system for transmitting the calculations of the computer as a series of signals to the adjusting apparatus. Let us now suppose that, were the gun fired in accordance with the calculations of the computer on a given occasion, the target would be hit; but let us also suppose that the signals transmitting these calculations have disturbing effects (whether on the adjusting

> apparatus or on the target) for which the computer has made no allowance. Accordingly, although the gun is set and fired in accordance with calculations that were correct at the time they were made, it nevertheless fails to hit the target because of the changes introduced by the process of transmitting those calculations.[4]

This example also conforms to our presystematic requirements concerning what shall be included in the extension of the term 'self-vitiating'. In this example it is the utterance of the calculations which are causally efficacious in bringing about the result. And if there are "essential" differences between this example (and others of its kind) and those indicated in Chapter I, these differences have not yet been clearly delineated.

Returning to the arguments of John Venn we see that either his arguments are too weak or too strong. They are too weak if he assumes that self-vitiating is a phenomenon which concerns the social scientist only. But if he is claiming that the phenomenon of self-vitiation makes successful prediction of the future impossible, then his arguments are too strong. For, as we have seen, examples of self-vitiation may occur in a branch of science as sacrosanct as that of celestial mechanics.

---

[4] op. cit., p. 469

"Predictions of Future Knowledge and Knowledge of the Future" - The Arguments of Popper

The concept of scientific methodology which is adopted for the purposes of this essay is essentially that which is advocated by Karl Raymond Popper. The requirements for what constitutes a scientific explanation or prediction are in most respects due to the early work of this man. What we have called "the Hempel-Oppenheim thesis" is sometimes referred to as "the Popper-Hempelian thesis".[5] It therefore is somewhat anomolous to have to deal with his arguments for the thesis that "no society can predict, scientifically, its own future states of knowledge".[6] This claim seems to conflict with his frequent assertions to the effect that the methodology in the social sciences is identical to that of the physical sciences:

---

[5] S. A. Grunbaum, op. cit.

[6] Popper's arguments are scattered throughout three major publications: The Poverty of Historicism, (Boston: The Beacon Press, c.d., 1957); The Open Society and its Enemies (London: Routledge & Kegan Paul, Ltd., 1945); The Logic of Scientific Discovery (London: Hutchinson, 1959). The above quotation is from The Poverty of Historicism, p. xi.

> From the point of view of scientific method, a social hypothesis suggested by self-intuition is no different position from a physical hypothesis about atoms. The latter may also be suggested to the physicist by a kind of intuition about what atoms are like. And in both cases, this intuition is a private affair of the man who proposes the hypothesis. What is 'public', and important for science, is merely the question whether the hypotheses could be tested by experience, and whether they stood up to tests.
>
> From this point of view, social theories are no more 'subjective' than physical ones.[7]

Popper does not believe, however, that this point of view is cogent for all social science. That is to say, history, and predictions about the future course of human history are excluded. Since Popper gives his own summary of his argument, we may perhaps be pardoned for allowing him to speak for himself:

> I have shown that, for strictly logical reasons, it is impossible for us to predict the future course of history...
>
> I propose to give here, in a few words, an outline of the refutation of historicism. The argument may be summed up in five statements, as follows:
>
> (1) The course of human history is strongly influenced by the growth of human knowledge. (The truth of this premise must be admitted even by those who see in our ideas, including our scientific ideas, merely the by-products of material development of some kind or other.)

---

[7] The Open Society, p. 308, n 14.

(2) We cannot predict, by rational or scientific methods, the future growth of our scientific knowledge. (This assertion can be logically proved, by considerations which are sketched below.)

(3) This means that we must reject the possibility of a theoretical history; that is to say, of a historical social science that would correspond to theoretical physics. There can be no scientific theory of historical development serving as a basis for historical prediction.

(4) The fundamental aim of historicist methods... is therefore misconceived; and historicism collapses.

The argument does not, of course, refute the possibility of every kind of social prediction; on the contrary, it is perfectly compatible with the possibility of testing social theories - for example, economic theories - by way of predicting that certain developments will take place under certain conditions...

The decisive step in this argument is statement (2). I think that it is convincing in itself: if there is such a thing as growing human knowledge, then we cannot anticipate to-day what we shall know only tomorrow.[8]

Before examining this argument in detail we may show that it is directly related to the problem we are considering in this essay. One of the crucial claims has to do with the assertion "We cannot anticipate today

---

[8] The Poverty of Historicism, pp. ix - x.

what we shall know only tomorrow." What Popper apparently has in mind is something like the following: suppose that we predict that in the future we shall know a general law L. But if we can formulate L, this means that our knowledge of L is not future buty present. Therefore we cannot predict that our knowledge of L will be limited to the future. In our terms such a fore-diction would be self-vitiating. Such fore-dictions may be treated as a small sub-class of self-vitiating hypotheses.

Popper, however, also deals with the more general phenomena of self-vitiation in his examination of what he calls the "Oedipus effect":

> The idea that a prediction may have influence upon the predicted event is a very old one. Oedipus, in the legend, killed his father whom he had never seen before; and this was the direct result of the prophecy which had caused his father to abandon him. This is why I suggest the name 'Oedipus effect' for the influence of the prediction upon the predicted event...whether this influence tends to bring about the predicted event, or whether it tends to prevent it.
>
> Historicists have recently pointed out that this kind of influence may be relevant to the social sciences; that it may increase the difficulty of making exact predictions and endanger their objectivity. They say that absurd consequences would follow from the assumption that the social sciences could ever be so far developed as to permit precise scientific forecasts of every kind of social fact and event, and that this assumption can therefore be refuted on purely logical grounds. For, if such a novel

> kind of scientific social calendar were constructed and became known...it would certainly cause actions which would upset its predictions.[9]

Just how much of the preceding paragraph Popper himself will subscribe to is by no means clear. Later statements suggest that he subscribes to the thesis that "an exact scientific calendar of social events, comparable to, say, the Nautical Almanac, has been shown... to be logically impossible".[10]

A Critique of Popper's Views

There are a number of difficulties involved in Popper's arguments which need to be unravelled prior to evaluating the cogency of these arguments.

First of all, the indicator term 'we', which appears in his statement (2) above, is not particularly relevant to the argument.[11] The point at issue is whether or not a rational being could predict the future course of human knowledge and hence the future course of human history dependent upon that knowledge. When

---

[9] Ibid. p. 13

[10] Ibid. p. 37

[11] We are not criticizing Popper's attack on historicism with which we are in full agreement; but only his claim that it is logically impossible to predict the future course of human history.

the question is thus rephrased much of the puzzle which Popper sees in the issue simply disappears. In fact it seems that his argument is in some respects based on a tautology: "Future knowledge is not present knowledge, and conversely." In this age of space exploration, our imagination concerning "possibility" (however construed) need not be limited to human knowledge of human knowledge. There is surely nothing logically impossible in the assumption that somewhere in the vast reaches of the Milky Way there exists some rational being or race of rational beings who, with a knowledge of a thousand civilizations on as many planets, could formulate a theory which, given the antecedent conditions of human civilization up to the present, could with some high degree of accuracy predict the future course of human knowledge. Popper's logical impossibility thus turns out to be a phantom. In fact all his argument seems to say which is tenable is "For any human being **A**, if **A** does not know **L**, then **A** does not know **L**." And this as we noted above is simply a tautology. Hence part of the puzzle disappears if we remove the indicator term 'we' from Popper's arguments.

Then the question as to whether or not we know now what we do not know now is seen to demand a negative

reply. But this reformulation, rather than enhancing his argument also simply renders it innocuous. And we cannot infer from this argument that therefore the future course of human knowledge is unpredictable due to a logical difficulty.

Third, whether or not any event can be successfully predicted depends in large measure on the body of theory from which the statements in the explanans are derived. The same argument holds in regard to predictions of the future course of human knowledge. Whether or not human knowledge is predictable will depend in large measure on whether or not there are sufficient law-statements which are highly confirmed and which, taken together with certain statements describing relevant antecedent conditions will logically entail that in the future a particular piece of knowledge will be had by human beings. On the basis of our present understanding of human and natural history it appears plausible to assert that the first atomic pile was produced in the twentieth century. Now according to Popper's thesis it would have been impossible for any scientist to predict, say, prior to 1850, that by mid-twentieth century we would a) know certain properties of the uranium atom, or b) know how to produce atomic fission, or c) that President

Truman would decide to use this power to destroy two Japanese cities. Thus not only is prediction of human knowledge impossible, according to Popper; the same will hold for future history. The question, however, is, given evidence of a number of other civilizations of rational beings plus a creative imagination, would it be logically possible for a rational being to construct a theory which could predict the future course of human history? There seems to be no legitimate reason for a negative reply to this question. If the results of the prediction were made public would the hypotheses on which it was based prove to be self-vitiating? The answer to this question will in large part depend on the kind and specificity of the event being predicted.[12]

Suppose that a rational being with a body of sociological and psychological theory and evidence concerning human civilization were to predict that human civilization would possess a body of laws L_1 ... L_n not prior to the year 2000 A.D. Suppose further that this

---

[12]We shall have occasion to note that Popper's thesis fails to make clear the highly ambiguous character of the term 'knowledge'. And until this ambiguity is removed it is difficult to assess clearly his claim regarding the inability of any society to predict, scientifically its own future states of knowledge.

fore-diction becomes known to the scientific community in 1963. Now we may ask, "Under these circumstances, would the fore-diction be self-vitiating?" The answer to this question will depend on the kind of designators that are used in the prediction to refer to $\underline{L}_1 \ldots \underline{L}_n$. Do '$\underline{L}_1$' ...'$\underline{L}_n$' in the prediction simply name the laws (e.g. "Boyle's law", "Newton's laws of motion")? Or do they characterize the laws (e.g. "laws regarding the rate of disintegration of the uranium atom", "laws of the refraction of light waves")? Or do they explicitly state the laws (e.g. "$E = MC^2$", "Every body at rest or moving in uniform rectilinear motion continues at rest or in uniform rectilinear motion until acted upon by some external force.")? If the prediction simply names the laws or characterizes them, the self-vitiating effect probably would not appear. Even if the laws are explicitly stated it is not clear that even in this case the fore-diction would be self-vitiating. Whether or not it would be self-vitiating would depend on the kind and degree of understanding which existed on the part of the scientific community in regard to what the laws indicated.

This latter point may be made clear by a somewhat simple-minded example. We may predict in regard to some child A, that on A's twenty-first birthday, and not before,

he will know the law "Ontogeny recapitulates philogeny". If the child become acquainted with tokens of the fore-diction we may ask whether the fore-diction was in his case self-vitiating. The answer will, it is clear, depend on what is meant by the term 'know'. At this point we can distinguish several ways in which the term 'know' can be understood in this context.

1. The term means merely that the individual is acquainted with (i.e. has had a sensory experience of) a certain set of tokens.[13] If this is the meaning of the term in the illustration, then clearly the fore-diction is self-vitiating as soon as the child hears the tokens produced by the predictor.

2. The term means that upon hearing or seeing the tokens, one will be able to reproduce replicas of the tokens or reasonable facsimiles thereof. If this is the sense of the term being employed, then whether the fore-diction is self-vitiating will depend on whether the child is able, because of having heard the fore-diction, to repeat the words 'Ontogeny recapitualtes philogeny' prior

---

[13]Later we shall need to distinguish between tokens and the names of tokens. This distinction shall prove important for an understanding of some of the issues involved in the problem before us. At this stage of the analysis we will assume that the individual is acquainted with the tokens rather than the names of the tokens.

to his twenty-first birthday, and in such a way that the tokens he produces are recognizable replicas of the original. If, after hearing the fore-diction, **A** is able to reproduce the tokens in such a manner we may say that the fore-diction was self-vitiating.

3. The individual, on or after his twenty-first birthday, is acquainted with the semantical rules of the language in which the fore-diction is made; and consequently he would be able to use the individual terms of the fore-diction in other statements according to the semantical rules of the language. In our example the fore-diction would not be self-vitiating unless **A** were extremely precocious.

4. The individual, on or after his twenty-first birthday, is provided with a body of evidence sufficient to confirm the law statement in question. If this is what the fore-diction intends by the term 'know' then most likely **A** does not know the law "Ontogeny recapitualtes philogeny" when he first hears the fore-diction. Hence the hypothesis is not self-vitiating.

5. The individual, on or after his twenty-first birthday, is acquainted with the scientific procedures and the kinds of evidence required to confirm or disconfirm the law statement in question. Again, the fore-diction

would not be self-vitiating if the term 'know' is construed in this way.

In regard to 4. it may be remarked that we have various ways of confirming particular beliefs. Or put in a slightly different way, there are a number of grounds on which we warrant our acceptance of certain statements. In common parlance we are often said to <u>know</u> certain statements on the basis of some authority. Most of our knowledge about the world is of this character. I may be said to know that Castro is the present leader of the island of Cuba, that the Chinese mainland has a communist government, that John Glenn was the first American to circle the globe in a space capsule. In this sense of 'know' I am not directly acquainted with (i.e. have not had any sensory experience of) any body of evidence which would <u>directly</u> confirm or disconfirm the statements in question. Nor am I clearly conscious of the procedures which would be required to confirm or disconfirm the statements. I know in the sense that an authority has informed me. This is not to suggest that the authority is beyond challenge. He may be mistaken, deceitful, deluded. And hence his statements may be false. Yet the point is that in this sense of 'know' the authority remains unchallenged; and I may or may not be acquainted with evidence or procedures which would tend to confirm or disconfirm his assertions.

The interesting case occurs when the predictor may be such an authority. Whether or not the fore-diction in such a case would be self-vitiating (in the sense of 4.) will be determined by the attitudes (i.e. the rationality) of A and also his conviction of the reliability, in the sense just indicated, of the person making the prediction or uttering the fore-diction. If the present attitude is one of rational trust in an authority, then the fore-diction in this sense will be self-vitiating.

These several senses of the term 'to know' make Popper's claim quite uncompelling in the absence of any precise specification by him of a sense in which predictions about the future course of human knowledge are impossible. As we have seen, whether or not such a prediction contains a self-vitiating explanandum will depend on which sense of 'know' is intended. And it has been shown that there are a number of senses in which the term may be employed in predicting future knowledge which do not have the self-vitiating effect.

These five meanings of the term, however, are by no means exhaustive of the possible ways in which it is used. The following schema is meant to be suggestive of some of the other senses in which the term 'know' can be understood. This schema is based on the distinction

between knowing *that* and knowing *how to*.[14] This schema will also bring to light the extremely ambiguous character of Popper's claim that we cannot predict future knowledge.

6. *X* predicts that at time *t*, *Y* will know the statement *S*.

7. *X* predicts that at time *t*, *Y* will know that the statement *S* is true.

8. *X* predicts that at time *t*, *Y* will know how to do *Z*.

A number of other examples can be treated as special cases of 6 - 8. A special case of 6. is:

9. *X* predicts that at time *t*, *Y* will know the statement "S_1 is true".

Another special case of 6. which shares certain features of 8. is:

10. *X* predicts that at time *t*, *Y* will know the statement "At time t_1, *Y* will know how to do *Z*".

A special case of 7. which shares features of 8. is:

11. *X* predicts that at time *t*, *Y* will know that the statement "At time t_1, *Y* will know how to do *Z*" is true.

---

[14] The distinction is made by Gilbert Ryle, *The Concept of Mind*, (New York: Barnes and Noble, Inc., c.d., 1949) pp. 25-61, we hold no brief for the distinction which may prove to be untenable. Our concern is merely to point out some of the ambiguity in the expression 'to know' and to point out also how this factor is of considerable importance for the claim that knowledge of our future knowledge is impossible.

The preceding schema enables us to produce examples of ever-increasing complexity. By re-writing variables and combining elements of 6. and 11. we can get:

> 12. Z predicts that at time t, W will know the statements "X predicts that at time t_1, Y will know that the statement 'At time t_2, Y will know how to do Z' is true".

In this example, the verb 'to know' is used in a variety of ways. The first occurrence is subject to the demand for clarification in terms of 1. - 5. above. That is to say, it may mean merely that W will be acquainted with a set of tokens which are replicas of those which follow the term 'statement' (first occurrence) in 12. Or it may mean that W is acquainted with a body of evidence which will tend to confirm or disconfirm the statement (i.e. the expression beginning with '"X predicts..."') embedded in the illustration. Or it may be taken in one of the other ways in which the senses of the term 'to know' is delineated in 1. - 5.

The second occurrence of the term 'know' in the above example seems to accord more closely with the sense characterized by 4. and 5. The third occurrence of the term is used in the sense of 8. But even this occurrence is not without ambiguity. Consider, for example, the locution, "X knows how to play the piano." In the most common sense of this term it means that X has the ability to play the piano. But a number of qualifications are needed to make clear what is intended. X may have broken

some of his fingers or one of his arms. Or his hands may be crippled with arthritis. What the locution then intends is the counterfactual, "If X were not crippled by arthritis (or suffering from broken fingers or arms) then he would be able to play the piano".

The need for other qualifications is equally apparent. A composer, for example, himself a violinist, may know how to play the piano in the sense that he is acquainted with the rules of musical composition and can discriminate between playing a piece of piano music properly and improperly according to the rules of accepted piano technique. He may also have acquaintance with the various names of the notes of the piano-forte and be acquainted with the rules which prescribe which fingers to use in striking certain combinations of keys, yet he himself be totally unable to play the piano,

On the basis of these considerations it is readily seen that 12. is filled with ambiguity. And whether any such fore-diction is or is not self-vitiating will depend first of all on specifying the exact sense in which the hypothesis was intended. Until this is done, the question of whether or not such an hypothesis is self-vitiating cannot be answered.

Similar arguments apply to the other forms 6 - 11. A teacher, on the basis of past experience and certain general laws of learning, may successfully predict when a given student will have mastered a particular piece of music.

Likewise, on the basis of past experience, a teacher may be able to predict that a given student will know at time t that the statement S is true, whereas formerly the student was acquainted with tokens of S but was not aware that S is true and may even have believed that S is false.

What is evident from the fore-going analysis of Popper's claim is that certain fore-dictions of human knowledge but not every such fore-diction, will be self-vitiating. But whether any particular fore-diction is in fact self-vitiating will depend on a number of factors including the question of what was intended by the fore-diction.

Even more important than the above considerations is the question whether there are logically-unavoidable cases of self-vitiation. It can be shown that the kinds of cases we have been considering offer a number of possibilities for avoiding the self-vitiating effect. Venn suggests a variety of ways when he says, "The publication of any prediction about the conduct of human beings (unless it were kept out of their sight, or expressed in unintelligible language) almost certainly would have some effect".[15] A fore-diction S could be uttered in a foreign language which was not understood by the person(s) about whom the prediction was being made. This move, in effect, prevents

---

[15] Op. cit., p. 576

such a person or persons from knowing the prediction except perhaps in the sense of 1. and possibly in the sense of 2. In the sense of 3 - 5 however, since the prediction is made in an unfamiliar language and presumably the persons about whom the prediction is made are unfamiliar with the semantical rules of the language, the self-vitiating feature is effectively removed. But it is possible, surely, that if the fore-diction were uttered in English and concerned the linguistic behaviour of a person who spoke and understood only the French language, then the fore-diction would not be self-vitiating even in the sense of 1. and 2. This claim can be made clear by way of an example. I may predict that a French-speaking child B will be acquainted for the first time with the French translation of the law-statment "Ontogeny recapitualtes philogeny" on his twenty-first birthday. Even though B is presumably acquainted with the tokens contained in my fore-diction the hypothesis is not vitiated. For what the hypothesis refers to are certain tokens in the French language rather than the English.

If 'know' is taken in the sense of 1. or 2. there are other similar ways in which we can predict events in the manner of 6. and at the same time avoid the self-vitiating character of the fore-diction. We may, for example, instead of predicting that at time t, Y will know the statement "Ontogeny recapitualtes philogeny",

use the elliptical form:

> At time t, Y will know that the statement indicated by three words: the first of which is composed of the fifteenth, fourteenth, twentieth, fifteenth, seventh, fifth, fourteenth and twenty-fifth letters of the English alphabet; the second of which is composed of the eighteenth, fifth, third, first, sixteenth, ninth, twentieth, twentyfirst, twelfth, first, twentieth, fifth, and nineteenth letters of the English alphabet; and the third of which is composed of the sixteenth, eighth, ninth, twelfth, fifteenth, seventh, fifth, fourteenth, and twenty-fifth letters of the English alphabet.

Here again the particular difficulty of self-vitiation is obviated in one sense; although in another sense it may not be. Whether or not the individual, Y, will be motivated to look up the words in question in a standard dictionary will be a matter of some concern presumably to the psychologist. We have shown, however, that no inherent logical difficulties remain in this formulation of the fore-diction. As long as the prediction concerns Y's acquaintance with tokens, we can always find suitable means for avoiding the use of those tokens; i.e., we can name the constituents of the tokens or use definite descriptions or other descriptive or naming devices to designate the tokens referred to in the fore-diction.[16]

---

[16] Whether such devices are acceptable options for preventing the self-vitiating effects in a society which values the free interchange and exchange of information is a highly debatable question. And in our opinion it is not debated with sufficient frequency. Merton writes, "To be sure, in misanthropic mood, or in excessive devotion to the values of social science above all other human values, or in the self-defined role of a scientific samurai,

These remarks are not meant to indicate that certain fore-dictions may not have causal consequences such that the events they predict are not borne out. Indeed many (if not all) self-vitiating hypotheses turn out to have precisely this characteristic. All we have been concerned to show up to this point is that Popper's logical thesis is unsubstantiated.

The Arguments of Feuer and Gewirth

Others have made similar claims and have gone on to argue for the difficulties or impossibilities of making predictions about the future in the social sciences. Lewis S. Feuer has argued for this position in the field of economics.

> What the directions of economic development are, whether toward stagnation, collapse, or prosperity, depends primarily upon the emergence of new industries. After the laws of economics

---

the social scientist might write out, seal and safely deposit his prediction of impending unemployment or war or internecine conflict, bringing it to light only after the predicted events had come to pass. But this would be almost as reckless of the body politic as of his own corporeal self. When one considers the profound objection of many individuals to being regarded as psychological guinea pigs, one can roughly imagine the aggregated fury of an entire population upon discovering itself transmogrified into one immense sociological guinea pig." Op Cit., p. 130. The conditions under which an entire population would react in the manner described by Merton are not altogether obvious. But one thing appears quite clear: the kinds of alternative solutions to the problem of self-vitiation we have listed raise questions of value of considerable importance to a "free" society.

> have been formulated, laws concerning prices, wages, the rate of interest, it is still not possible to foretell what the human intelligence will discover during this stage of economic development. There was no forecast from any economic school that the fission of the uranium atom would be discovered, and that a new mode of industrial energy would become available.[17]

While Feuer's comments are meant to apply principally to capitalism as an economic system they are relevant for problems of wider scope. They seem to represent a widespread conviction that human creativity is unpredictable in principle.

> Creative response, as a social phenomenon then, introduces variables which do not fall within a determinate set of economic equations. A surd factor enters into the course of technical evolution. The directions of economic development therefore cannot be predicted from the principles of economic theory.[18]

He concludes with the following paragraph:

> Most important, there is no way of foreseeing whether a given era will be marked by the dearth or the emergence of new great industries. It remains indeterminate therefore whether investment waves will be forthcoming which would enable the capitalist system at any given period to surmount the pressures of underconsumptionist disequilibrium.[19]

---

[17]Lewis S. Feuer, "Indeterminacy and Economic Development", Philosophy of Science, 15: 234, July, 1948.

[18]Ibid., p. 235.

[19]Ibid., p. 241.

These arguments somewhat parallel Popper's claim concerning the logical impossibility of predicting future knowledge. Feuer, however, simply claims that we cannot predict what creative responses will be forthcoming in the future. But this claim seems to be based on a non sequitur. The fact that no economist did in fact predict the fission of the uranium atom is hardly an argument that such predictions are impossible in principle. More generally, the fact that we, at present, have no body of theory which would enable us to predict so-called creative responses is hardly a sufficient argument to substantiate the claim that no such body of theory is possible.[20] One of the crucial questions which any such claim must answer is, "What is meant by the term 'possible'?". For if it is logical possibility which is intended then, as we have seen, the claim is untenable. If it is empirical possibility (or impossibility) which is intended, then the claim that a particular activity or piece of knowledge is beyond the

---

[20] Often in the literature the position here described offers us more rhetoric than argument. Charles Hartshorne, for example, writes,

> "The idea that to know is to be able to predict, so that any inability to predict must mean partial ignorance, is, I hold, untenable. The ultimate function of knowledge is not to foresee, but to create. The two are by no means coincident. We predict most perfectly astronomical phenomena where we have no creative control. We control most perfectly our bodily behavior, but how vague and uncertain are our predictions of this behavior!...

"possibility" of science or human reason is perhaps only to baptize our ignorance, or to give evidence of an infertile imagination, or to exhort us to stop trying to reach "beyond our grasp". Baptism of ignorance is hardly a noble philosophical pursuit; and claims that our reason is limited in such and such a way are usually short-lived or quickly disproven by new discoveries. We may readily admit that the human mind is so constituted that it cannot know everything. But it does not appear to be particularly warranted to claim that one has discovered what the precise limitations of the human mind happen in fact to be.[21]

> "Is the goal of psychology to be able to write the poet's poem beforehand by predicting it, to compose the composer's music by telling him what it is going to be? Should the psychologist be able to foresee, and thus make the creative decisions, and their expression in speeches and laws, by which statesmen, or an entire people resolve some political tangle? How extravagant can the theory of science become? I suggest that, since creative activity is that which leaps unpredictably from its causal base, the proper function of psychology in regard to creation is to increase our freedom from the compulsions and fixations which often make behavior only too largely predictable."

"Freedom Requires Indeterminism and Universal Causality." Journal of Philosophy, 55:800-801, September 11, 1958.

[21] The kind of arguments we have been considering seem to be motivated by an inordinate fear of scientists' being able to predict the creative response of human individuals. It is as though if we were able to predict such responses we would automatically dry up the wellsprings of creative endeavor. If we were to respond to rhetoric with rhetoric we might reply in this way:

> "How many Beethovens and Mozarts have gone to their graves because no one taught them to sing

In a later article, Feuer makes it clear that the kind of possibility he has in mind is not that of logical possibility:

> A directional law of evolution is not logically impossible. The law of entropy in physical science asserts that for closed systems there is a definite direction that energy transformations will take. A directional law of social change cannot therefore be excluded on purely logical grounds. The law of entropy, however, holds only for closed systems; if there were an influx of energy from other parts of an unlimited universe, a given region might not actually evolve toward thermodynamical equilibrium. And laws of history are beyond our grasp precisely because human initiative, human intervention, upsets any approximation toward a closed social universe. The factor of human creativity corresponds to the influx of energy into the physical system. It makes moreover for the unpredictable aspect of historical causation.[22]

Unless we are misconstruing Feuer's comments it appears that the phrase "closed social system" is functioning as a metaphor in the argument. The same appears to be true in regard to the use of the phrase "law of entropy"

> the melodies ringing in their souls? And how many Newtons and Einsteins have perished without giving to the world the answers to some of the burning problems of mankind? And how many children have died of cancer because a Harvey or a Lister or a Pasteur died without performing the experiment that was in him?
>
> "And how many of these perished, these unknown untutored geniuses of art and science, because no one could look into their souls and predict that here is a genius who will contribute immeasurably to the welfare and well-being of mankind in the following specific ways."

[22]Lewis S. Feuer, "Causality in the Social Sciences", Journal of Philosophy, 51: 694, November 11, 1954.

when it is applied to a closed social system. If we are not misconstruing his arguments and these terms indeed function metaphorically then I must admit that the force of his objections escapes me. If, however, Feuer means us to take these comments as empirical statements about matters of fact, then as far as I am aware, his arguments find little if any support from the physical or the social sciences. In fact, prima facie it appears that his claim that "human intervention upsets an approximation toward a closed social universe" may indeed be false.[23] At any rate, Feuer has, it seems, failed to justify his thesis that human creativity makes "for the unpredictable aspect of historical causation."

Alan Gewirth also has argued that certain foredictions of social events can be vitiated by the creative response of the individuals about whom the prediction is made. He treats self-vitiating hypotheses as a sub-class of phenomena which he calls the "reflexive action of men on social laws". He writes:

> Now the general position I wish to present is the following. In dealing with social phenomena, social science deals largely if not entirely with things which impinge directly on men's values - wealth, power, various kinds of interpersonal relations, and so on. The aim of social science may be said to be to attain knowledge of the laws

---

[23] For an interesting and illuminating discussion of the concept of entropy in the physical sciences, see A. Grunbaum, op. cit.

> of these matters - that is, of their cause-effect relations. Since, however, man as conscious voluntary agent is in large part both the knower and the subject-matter of these laws, his knowledge of their impact on his values may lead him to react on the laws reflexively in order to change them. Consequently, the laws of the social sciences cannot have the same fixity or permanence as the laws of the natural sciences.
>
> This reflexive reaction of men on social laws has interesting logical as well as social consequences. One of its familiar aspects is found in such reflexive situations as the "self-destroying prophecy."[24]

In a footnote he attempts to give two formal characterizations of this type of prediction:

> 1. If there is a prediction P that a certain kind of action A will result in a certain consequence C, then this prediction has the result that the contrary...of action A occurs and that it results in the occurrence of the contrary of C.
>
> 2. If there is a prediction P that a certain kind of action A_1 will result in a certain consequence C, then this prediction has the result that action A_1 and also other action A_2 occur, (sic) and that these actions have the result that the contrary of consequence C occurs.[25]

In Chapter VI we shall have occasion to examine his arguement in more detail. At this point we can remark two difficulties which appear, at first glance, to render his formalization inadequate. First it is by no means clear what is meant by the phrase 'the contrary of an action'.

---

[24] Alan Gewirth, "Can Men Change Laws of Social Science?", Philosophy of Science, 21: 230, July, 1954.

[25] Ibid.

Perhaps this is simply an elliptical way of saying "the contrary of the statement describing action A". If our rephrasing meets Gewirth's intent, then another puzzle appears: what can be meant by the term 'the contrary' in this case? For example, what would constitute the contrary of a sentential function 'Fx'? Presumably a contrary would be one which is a universal generalization of this function. In other words the function universally generalized would be '(x)Fx'. And a contrary of this statement would be '(x)~Fx'. But Gewirth's own symbolization prevents this interpretation, since he uses an existential quantifier in such a way as to preclude it.[26] The opacity of his symbolism, which accords with no logistic system with which this author is acquainted, makes his intent impossible to fathom.

Second, he in no way specifies what are to be the values of the variable 'A_2'. Unless the range of values of this variable is clearly specified we are in no position to discriminate between those predictions whose explanans contains faulty laws or unreliable evidence and whose explanandum, although in no way causally related to the

---

[26] Gewirth's symbolism of 1. and 2. above is as follows:

1. $(\exists x)(Px[A \rightarrow C]) \longrightarrow \{(\exists y)(\sim Ay) \longrightarrow (\exists z)(\sim Cz)\}$

2. $(\exists x)(Px\,[A \rightarrow C]) \rightarrow \{[(\exists y)(A_1 y)\ .\ (\exists z)(A_2 z)] \rightarrow (\exists w)(\sim Cw)\}$

The double tilde is meant to represent, not the negation, but the contrary of the expression(?) which follows.

subsequent turn of events, turns out to be false; and those self-vitiating predictions whose explanans contains hypotheses which we would accept as true, at least prior to the demonstration of the self-vitiating character of the explanandum.[27]

The body of Gewirth's article, however, attempts to show that self-vitiating fore-dictions illustrate man's ability to change social laws; and if Gewirth's arguments are persuasive, we shall have to abondon the empiricist claim that no demonstrable differences have been shown to obtain between the physical and the social sciences. This possibility is hardly at issue, however, inasmuch as he seems merely to be proposing a new usage of the term 'social law'.

> Let us consider three different relations which man may bear to scientific laws: namely, that he can <u>exchange</u> the laws, <u>apply</u> the laws and <u>create</u> the laws. Man <u>exchanges</u> laws when he replaces the conditions under which one set of laws - i.e., one correlation of variables - obtains, for another set of conditions under which a different correlation obtains. Man <u>applies</u> laws when he uses a correlation

---

[27]Somewhat later we shall find it necessary to modify the fourth requirement of the Hempel-Oppenheim thesis for our purposes. Their requirement is that the statements in the explanans be true. Now this requirement may indeed be retained as a criterion for determining what constitutes a scientific explanation. For our purposes, we need to weaken the criterion to admit of those explanantia which contain statements which we would accept as true prior to the determination of their self-vitiating character. It should be noted that we are not advocating a weakening of the criterion for a scientific explanation, but only of the criterion for what shall constitute a self-vitiating fore-diction. The reason for weakening this requirement is simply this: if we retained the stronger criterion, the definition which we shall propose would turn out to have a null extension.

> which exists independently of his decisions to produce an effect which would not have occurred if he had not made use of this correlation. Man creates laws, finally, when by means of his free decisions and consequent action, he causes a correlation to exist which did not exist before.[28]

Gewirth's error arises out of his failure to realize the tenselessness and non-existential import of a law-like "correlation". It is therefore likely that the distinction which he wishes to draw cannot be maintained according to his criteria. Two examples will illustrate this contention:

Assuming that the first self-propelled heavier-than-air craft was flown in the first decade of the twentieth century we may claim that in this instance man exchanged laws. In this instance man replaced the conditions under which one set of laws obtained, i.e., the laws of gravity, for another set of conditions under which, in addition, a different correlation obtains, namely, the conditions described by the laws of aerodynamics. But this example also conforms to Gewirth's description of what it means to apply a law; i.e., in this instance men used a correlation which exists independently of any decision on their part to fly an airplane. What Gewirth wishes to maintain is that in social science phenomena what often happens is that man "causes a correlation to exist which did not exist before" whereas in physical phenomena (such as that in the example of the airplane) the correlation of

---

[28] Op. cit. p. 234.

variables is not created but existed prior to its discovery. Causing an instance of such a correlation is what Gewirth means by creating laws. Yet we may inquire as to whether or not the flight of the first airplane cannot also be subsumed under this classification. But perhaps the matter can be put more illuminatingly.

On the basis of our assumption, no heavier-than-air craft had flown prior to the twentieth century. The problem then may be stated in the form of a counterfactual: "If someone had built a craft of the type known as a Piper Cub, prior to the twentieth century, and had been able to supply it with fuel, it would have flown". Now apparently Gewirth wishes to claim that we are warranted in making some such counterfactual claim in regard to natural laws but that we are not warranted in making them in regard to social laws. But if this is the case, then what we need are other criteria for distinguishing between natural and social laws. We cannot make the distinction on the basis that in a social law man creates a correlation which did not previously exist unless we have independent evidence for determining whether the correlation was confirmed -- e.g., evidence constituted of something other than the described human intervention. Otherwise we are in peril of doing one or the other of two things: we determine that the correlation did not exist previously because the law in question is a social law, and hence involve ourselves in a circular argument; or we treat the

assertion "If L is a social law, then the correlation of variables which obtains in L was not instantiated prior to human intervention." as analytic. In the latter case we have done little more than formulate a rule governing our use of terms; but we have hardly shown how to discriminate between social laws and natural laws or between correlations which are confirmable prior to specified kinds of human intervention and those which do not. If Gewirth's contention is to be maintained, what is required is a means for determining when a correlation of variables "exists prior to human intervention" and when such a correlation is created de novo. It is by no means clear how such a question could be answered and Gewirth's analysis offers us no guide. And inasmuch as the question appears unanswerable, Gewirth cannot be said to have made good his claim that in self-vitiating hypotheses we have a class of phenomena which enable us to distinguish between social and natural laws.

Gewirth makes a related claim concerning the logical relationships which obtain between the terms of social and natural laws:

> It is recognized that the relations of implication which hold between antecedent and consequent in natural laws are different from those which hold in the case of purely logical relations like those of mathematics. My point then is that the relations of implication in social laws are different from both of these. For both within the antecedent and between the antecedent and the consequent in social laws, human decisions do or may enter. Therefore, even the kind or degree of eternal truth found in natural laws

> which may be said to hold only for our present cosmic epoch, is not found in social laws.[29]

Whether or not we accept his enthymatic inference as valid will largely depend upon our willingness to accept his suppressed premise, viz. "If human decisions do or may enter within the antecedent or between antecedent and consequent of any law L, then the kind or degree of eternal truth found in natural laws will not be found in L". But this premise itself is plagued by the same obscurity which confounds other parts of his argument. What kind of entities are human decisions that they can or may enter the antecedent or between the antecedent and consequent of laws? Gewirth may have some kind of spatial or temporal entity in mind. For example, if someone propels a billiard ball toward another, human decision may prevent the first from striking the second by causing the placing of a hand or other object between them immediately after the initial impact. In such a case what we encounter is an example of a whole set of problems raised by our failure to make explicit the auxiliary assumptions regarding boundary conditions which are tacitly assumed in making a specific prediction. In the second half of Chapter IV we shall have occasion to examine this problem in greater detail. At this point it may suffice to note that our example of human intervention involved physical or natural laws rather than social laws. But if we were to accept Gewirth's classification, such

---

[29] Ibid., p. 237.

phenomena as that of moving billiard balls would need to be considered as social rather than as physical phenomena inasmuch as human decisions seem to be involved in upsetting predictions of the balls' behavior.

In this chapter we have examined what appear to be some of the representative arguments in the literature concerning the phenomena of self-vitiation. We have attempted to show that these arguments fail for a number of reasons not the least of which is the ambiguity and obscurity of the terminology employed in the arguments.

We shall now turn our attention to a more rigorous examination of the problem and attempt to explicate the notion of self-vitiation in a formal and more precise manner.

CHAPTER III

A FORMAL ANALYSIS OF THE PREDICATE 'SELF-VITIATING'

Introduction

The problem of self-vitiation involves two rather distinct features of hypotheses. It involves first of all a feature of those hypotheses which, due to certain characteristics, have already been demonstrated to be false. In this instance what we require is a manifest predicate which can be applied to the hypotheses in question.[1] On the other hand, there are those hypotheses which, on the grounds of their similarity to other manifestly self-vitiating hypotheses, we have reason to believe would be self-vitiating if they were uttered under certain conditions. What we require in this instance is a disposition predicate, i.e., a term which refers to a latent property or characteristic of hypotheses. The term which we shall employ in referring to the manifest characteristic of hypotheses is 'self-vitiating'. The term we shall employ in referring to the latent characteristic of hypotheses is 'self-vitiatable'. A formal definition of the manifest predicate 'self-vitiating' presents no particularly difficult obstacles. Defining its associated disposition predicate 'self-vitiatable', however, presents problems which are not easily resolvable. In addition to the general problems of defining

---

[1] We use the terms 'manifest' and 'disposition' according to the explication of Nelson Goodman. See Fact, Fiction and Forecast.

disposition terms, we face the difficulty that 'self-vitiatable' unlike such terms as 'magnetic' or 'soluble' is not found in any well-confirmed body of theory. Thus in attempting a formal definition of the dispositional predicate 'self-vitiatable' we face a double task: we must show not only how the term may be clearly defined in a way which will specify the range of application of the term; but also how the term may be incorporated into sociological, psychological and physical theory. The problems of defining 'self-vitiatable' will be treated in Chapter V. In the present chapter we shall confine ourselves to defining the manifest predicate 'self-vitiating' and problems related to this task.

First Steps

In this chapter we wish to mark off for investigation one particular class of phenomena which, until now, has been considered together with another important class. The distinguishing characteristic of this class of phenomena is that of a causal relationship which obtains between an utterance and the event described by the utterance. This formulation of the distinguishing characteristic is itself faulty as will become clear in a moment. What we wish to note at this point is that we are not considering that particular class of hypotheses which are better considered as a sub-class of the traditional paradoxes (syntactical and semantical). It would appear that the failure to mark this distinction between these two broad classes of

hypotheses (i.e. between those which are vitiated by entering into a causal relationship with the event described by the hypothesis, and those which are vitiated due to their self-referential or to their syntactical character) in the past has been a factor in preventing any clear and concise notion of what could be meant by 'self-vitiating', 'self-destroying', etc. By making the distinction as clear as we can, we may hope to mark off those areas of investigation which belong primarily to the social and physical scientists, on the one hand, and those areas which belong primarily to the logician, on the other.

The following symbols are introduced as formal aids in the analysis of the predicate 'self-vitiating'. Let

'h'	stand for	a particular utterance or an individual composed of a number of utterances containing n members (where 'n' is any positive integer) each of which contains as a sentential part a) a replica of the sentential parts of each of the others, or b) that which is logically equivalent to the sentential parts of each of the others, or c) that which may be construed as an adequate translation of the sentential parts of each of the others. And further that each such utterance constitutes a fore-diction that e.
'H'	stand for	the class of all the utterances which conform

		to the conditions specified under the description of h.[2]
'l'	stand for	an utterance which is the explanans of an explanation from which h is deducible. The requirements for h apply mutatis mutandis to the components of l.
'L'	stand for	the class of all utterances which conform to the requirements for membership in l.[2]
'x_1'...'x_n'	stand for	particular events in a causal sequence.
'Cx_ix_j'	stand for	'x_i causes x_j'.
'~e'	stand for	the event such that the substituend of '~e' is a replica of the negation of the sentential part(s) of the value of h.
'Iyh'	stand for	'y logically implies h'.
'Fh'	stand for	'h is false'.
'Vh'	stand for	'h is self-vitiating'.

'Self-vitiating' then may be contextually defined as indicated by the following schema:

$$Vh = \text{df. } (\exists x_1)(\exists x_2)\ldots(\exists x_n)(\exists H)(\exists L)(h \in H \,.\, (y)(y \in L \supset Iyh) \,.\, Chx_1 \,.\, Cx_1x_2 \,.\, \ldots \,.\, Cx_n{\sim}e \,.\, Fh)$$

Rendered in English this may be read: "To call a particular utterance or set of utterances h self-vitiating is to say that there exists a set of utterances L such that

---

[2] In a sense L and H are pseudo-classes.

every member of L logically implies h; that h is part of a causal sequence which eventuates in ~e; and that h is false."

A number of terms and conventions are adopted here which require some comment. The values of h are utterances which have sentential parts that are tokens. Utterances are, of course, events, where 'event' is here construed in its broadest sense; i.e., an event may be composed of a number of other events. Thus, for example, several members of H may be understood to compose a single spatially or temporally extended event. The single event so construed is termed an 'individual' in the sense of the calculus of individuals, i.e. as a sum of individuals.[3]

The term 'replica' is understood here as follows: "x is a replica of y if and only if x and y share certain specifiable physical characteristics". For example, 'DOG' is a replica of 'dog' as well as of 'D O G' insofar as these word-tokens are each composed of letter-tokens of the fourth, fifteenth, and seventh letters of the English

---

[3] Nelson Goodman, The Structure of Appearance, and Henry Leonard, Singular Terms, typescript, Widener Library, 1930. A later version by these two authors is found in "The Calculus of Individuals and its Uses", Journal of Symbolic Logic, 5:45-55, June, 1940.

alphabet in that order.[4] We have expressly avoided the introduction of the notion of "sign-types" in order to simplify the problem before us.[5]

We have added the phrase "or c) that which may be construed as an adequate translation of the sentential parts of each of the others;" in order to make clear what may not be clear by dealing solely with the notion of logical equivalence.[6] It is by no means always clear that an expression $\underline{S}_1$ is logically equivalent to $\underline{S}_2$ if the latter is an adequate translation of the former. Of course 'adequate translation' may be so defined that $\underline{S}_2$ is an adequate translation of $\underline{S}_1$ only if $\underline{S}_2$ is logically equivalent to $\underline{S}_1$. But, it would appear that this definition would be far too restrictive for our purposes. Suppose that one or more individuals make up $\underline{h}$, each of which contains tokens we would say are misspelled, syntactically incorrect according to the rules of sentence formation, or otherwise defective as an English sentence. The question as to whether this sentence (i.e. a sentential part of $\underline{h}$) was logically equivalent to some other English sentence

---

[4]We can extend the notion of "replicahood" to other senses by introducing the notion of an ideal observer: a set of phonemes $\underline{P}$ is a replica of an inscription $\underline{W}$ if an ideal observer $\underline{O}$ were to utter $\underline{P}$ upon being asked to read $\underline{W}$ aloud. Similar remarks obtain for braille inscriptions and sign-language.

[5]See Chapter II for the reasons for this decision.

[6]We shall later find it convenient to drop clause c). Reasons for this decision will be given presently.

which lacked the grammatical errors of the first, is beside the point. The question at issue would be whether or not the poorly constructed sentence entered a causal relationship with certain events in such a manner that the claim made by the utterance of which this set of tokens is the sentential part turns out to be false. Or put in a slightly different way, the question would be whether the receiver (whether the intended second party or not)[7] of these tokens could construe them in such a manner that he would understand them in a "similar enough" manner to the correctly formulated English sentence intended by the producer. At issue is whether or not such a poorly constructed sentence would contribute to the self-vitiating effect. If such an utterance would, under suitable circumstances produce the effect of self-vitiation, then it would appear that we need to include it in h even though the task of construing it as logically equivalent to the other individuals in h seems to strain the notion of "logical equivalence". But more of this point below.

'H' is to stand for the class of utterances which conform to the conditions specified under the characterization of h. We have used the class calculus purely for reasons of convenience. And we have assumed that it will

---

[7] See Henry Leonard, Principles of Right Reason, (Henry Holt and Co., New York), 1957, pp. 89 ff. for a discussion of sign-events. We shall use Professor Leonard's terminology except in those cases where we have reason to believe that a more nominalistic terminology will be helpful in our analysis. These latter circumstances will be indicated.

be possible to translate our definition (and any proofs dependent upon it) into the calculus of individuals. It may also be remarked that the way in which classes are introduced into the definition, no assumptions are made regarding what causal effects classes can have upon events. In other words, the ontology is left "open" in such a way that the definition would be serviceable in a system which admitted classes as values of its variables as well as in a system which admits only individuals as values of its variables.

We shall often wish to speak of the sentential parts of cerain utterances and of individuals composed of several utterances. It will be therefore at times convenient to speak of the sentential part(s) of h to write 'h (sp)'. This will make it clear that we are speaking of the tokens as physical concrete objects and not of the total sign-event (itself a complex bit of socio-psychological behavior) of which they are but parts.

The remarks concerning h and H apply equally to l and L. But in this case certain difficulties regarding logical implication become evident. Consider a case in which an individual A on the basis of a certain l* "deduces" h* which constitutes a fore-diction of e. We assume that, l*(sp) and h*(sp) have, however, one or more defects: either l* contains a syntactically incorrect formulation of an explanans which is to constitute part of an explanation; or h*(sp) is not logically deducible from l*(sp);

or h*(sp) is not syntactically correct. But the subsequent course of events warrants the counterfactual: 'If h* had not been uttered, e rather than ~e would have occurred.' Now we are faced with a choice. Either we can drop this kind of linguistic event from the extension of the term 'self-vitiating' and attend solely to those linguistic events which do not share these difficulties. Or we can lift some of the formal restrictions on the values of the variables. If we choose the first alternative, we are faced with the problem that we have so restricted the application of the term that its serviceability to the social scientist may be likewise restricted. If we accept the second alternative we seem to allow the notion of logical implication to be used in violation of certain logical laws, or in a sense which is at considerable variance from any well-established logical theory. If h*(sp) and l*(sp) are not well-formed expressions, then according to logical theory it is false to say that l* implies h*.[8]

These two alternatives are not exhaustive, however. The problem is to determine what may be construed as proper parts of h and l and whether or not mal-formed

---

[8]We shall use certain logical predicates in such a way as to allow them to apply indifferently either to utterances or to the sentential parts of those utterances. Hence we may write either 'h is false' or 'h(sp) is false'. In like manner we can write 'l implies h' or 'l(sp) implies h(sp)'. We may not, however, use this procedure with the predicate 'causes'. The reasons will presently be made clear.

expressions may be so construed. Another alternative would be to attempt to determine what an utterer, p, intended when he produced h*. We assume then that if the utterer p intended h when he said h* then h* is to be included in H. Likewise with l*. We could then re-write our definition in such a way as to preserve a rigorous sense to the term 'logical implication'. Logical implication would hold only between l and h and not between l* and h*. The question as to what an individual intends by his utterances is one which we assume can be explicated by psychological theory rather than by logical analysis of the tokens.

Then, under this interpretation, for any two sentences S_1 and S_2, S_2 is to be considered an adequate translation of S_1 if an only if a person p who utters S_1 intends to signify that which is signified by S_2.[9] If, then, it is theoretically possible to discover what p intends by an utterance S_1, then if S_1 is syntactically incorrect or otherwise defective, it will be possible to faithfully translate S_1 by S_2 such that S_2 is well-formed and signifies that which p intended by S_1.

The kind of utterance we are here considering, however, may still prove troublesome even if psychological theory is capable of explicating the sense of the expression

---

[9] We use the term 'signify' here in a sense similar to that of Professor Leonard. A statement S expresses the concern of the utterer p and indicates his topic of concern.

what p intends by S_1'. It may be the case that p is not consciously aware of what he intends by S_1, and that psychological theory is capable of demonstrating that fact. Then, if an utterance of S_1 should have some causal consequence such that S_1 turns out to be false, we would seem to be committed to including S_1 among the members of H. Presystematically, at least, S_1 seems to conform to the requirements of what would constitute a self-vitiating hypotheses.

None of the three alternatives proposed leaves us in a particularly felicitous situation. Later we shall propose a more adequate means for dealing with recalcitrant utterances such as h*. Before proceeding, however, we need to consider another type of difficulty which is closely related. This has to do with ambiguous expressions. In such a case we are uncertain as to whether a particular utterance ought to be included in H. In the case of an ambiguous expressions S_3, according to one translation S_3 will turn out to be self-vitiating, while according to another translation it will not be self-vitiating. An interesting example of this kind of situation is afforded by Shakespeare's account of the witches prophecy in 'Macbeth'. The "prediction" that Macbeth would never be slain by one of woman born, if interpreted as MacBeth interpreted it, could be construed as

self-vitiating.[10] Macbeth believed that the witches intended by this expression to predict that he would never be slain by a mortal. The belief in this "prediction" led Macbeth to act with an utter lack of caution for the safety of his own person; thus construed the "prediction" seems, in some presystematic sense, to turn out to be self-vitiating. If, on the other hand, the "prediction" is understood to mean that, as subsequent events make clear, Macbeth will die at the hand of one who "from his mother's womb was untimely ript", then the "prediction" on this interpretation (but, note, not as thus construed by Macbeth) is not self-vitiating.

Our initial problem which brought up all the above-mentioned difficulties involved specifying the criteria whereby we could determine whether an expression h* ought or ought not to be included in the extension of H. We were led to suggest that if h*(sp) could be adequately translated as h(sp) (i.e. what the utterer p intended by h* was h) then h* is to be included in the extension of H. It should be noted that, as we have been using 'adequate translation', the term 'logical implication' (i.e. the 'I' in our definition) may hold only between h and l and not between l* and h* even though the former are adequate translations of the latter. This particular outcome is not a happy one. For if the same criteria

---

[10] I am indebted to Mr. Peter Meloney of Michigan State University for this example.

are to be applied to the members of L, then the definition we have proposed for 'self-vitiating' will fail. It will fail due to the fact that there will be at least one member of L, namely, l* regarding which the relation of logical implication between explanans and explanandum will not hold: i.e. it is false that $(x)(x\epsilon L \supset Ixh)$.

The solution we propose to all of these difficulties is the following: we drop clause c) in its entirety and construe any utterance of h* or l* which are causally efficacious in effecting the event ~e as a constitutent of the value of some 'x_i' rather than as members of H and L respectively. In regard to this proposal we can retain our rigorous formal requirements concerning membership in H and L and at the same time avoid ignoring non-well-formed expressions which may be constituents of the causal sequence resulting in ~e.[11]

Any 'x_i' can take utterances or non-linguistic events as values. The expression 'x_i causes x_j' symbolized 'Cx_ix_j' is to be construed as meaning that x_j is described by a statement constituting an explanandum in a scientific explanation and x_i is referred to by an

---

[11] This criterion simply restricts the values of h to well-formed contingent hypotheses. We hold that any contingent hypothesis can serve as the explanandum of an explanation or prediction. That is, say, no event is inexplicable. By this we mean that at least the following statement is not self-contradictory: 'h is the explanandum of some explanation.'

essential constituent in an explanans from which the statement describing $\underline{x}_j$ is logically deducible.[12]

We are now able to deal with the recalcitrant utterances $\underline{h}^*$ and $\underline{l}^*$. We deal with them on the basis of an assumption: namely, that $\underline{h}^*$ and $\underline{l}^*$ are causally dependent upon some other pair of utterances $\underline{h}$ and $\underline{l}$. An elaboration of a previous illustration may help to make this point clear. Suppose the National Safety Council predicts that on a given holiday Hol. four hundred Americans will die in automobile accidents. Let us call this fore-diction '$\underline{h}_1$'. Subsequently a number of replicas of $\underline{h}_1$(sp) are published in newspapers and magazines; other replicas are contained in spot radio and television announcements. Let us call these utterances containing replicas of $\underline{h}_1$(sp) '$\underline{h}_2$' ... '$\underline{h}_n$'. As a result of hearing or reading these replicas and believing them to be true fore-dictions of the future, a number of individuals A, B, ... M are led to make remarks which are replicas of one or more of the following:

1. There's gonna be an awful lot of people killed this comin' weekend.

---

[12] Scientific prediction (or explanation) may be composed of various combinations of utterances:

a) All the utterances which are replicas of a particular utterance or set of utterances conforming to the H.O. requirements.
b) An individual composed of some of the members described in a).
c) An individual composed of some members described in a) plus some members of H.
d) any combination of a) - c).

2. I see by the papers that four hundred people are going to get it this weekend.

3. John said that he heard on T.V. that four hundred people are going to be killed over the holidays.

4. They're predicting another slaughter this weekend.

And so forth. Let us call replicas of these sentences '$\underline{h}^*_1$' ... '$\underline{h}^*_n$'. Although the various $\underline{h}^*_i$'s are in some sense "translations" of the $\underline{h}_i$'s, the adequacy of our definition is in no way dependent on such a construal. The term 'self-vitiating' can be applied to the members of $\underline{H}$ and withheld from the various $\underline{h}^*_i$'s. The $\underline{h}^*_i$'s are understood to contribute to the self-vitiating effect of $\underline{h}$; but are not construed as members of $\underline{H}$. In regard to our definition they are construed as being values of some '$\underline{x}_i$', i.e. they are essential constitutents of the causal sequence which begins with the utterance of $\underline{h}$ and eventuates in the event $\sim\underline{e}$. The question then as to whether or not the various $\underline{h}^*_i$'s are adequate translations of $\underline{h}$ is of no importance for our analysis. What is important is whether or not these $\underline{h}^*$ utterances are causally related to $\underline{h}$ and $\sim\underline{e}$. And the task of answering this question we suppose can be left to psychological investigation and some body of psychological theory.

While this proposal does solve the particular problems concerning the various $\underline{h}^*_i$'s, it does not appear to solve the problem of the logical relationship which holds between a particular utterance $\underline{h}$ and translations

of $\underline{h}$(sp) into certain foreign languages. Presumably the predictions of the National Safety Council are of sufficient importance to be included, in translation, in the various foreign-language publications in the United States. And presumably these translations also enter into the causal relationship which obtains between the fore-diction $\underline{h}$ and the outcome $\sim\underline{e}$. Ought we to attempt to show that an English language statement $\underline{h}_e$ logically implies a translation $\underline{h}_f$ into the French language? An affirmative answer to this question would lead us into difficulties with which we need not be concerned. One solution which is immediately apparent is that of treating all such translations of $\underline{h}_e$ as members of the class of $\underline{h}^*_i$'s. Such an alternative will not be open to us only in those situations in which predictions in English and predictions in some other language, say French, are deduced independently of each other and subsequently published in their respective languages. We may perhaps be justified in assuming that this kind of situation occurs only infrequently. And even when it does occur, it should pose no particular problem. In this case, if both the utterance in English, $\underline{h}_e$, and the utterance in French, $\underline{h}_f$, are required to bring about the self-vitiating effect, we can remove any ambiguity from the application of the predicate by saying that the individual composed of the utterances $\underline{h}_e$ and $\underline{h}_f$ is self-vitiating. Whether or not

either $\underline{h}_e$ or $\underline{h}_f$ taken by itself would be sufficient to bring about ~$\underline{e}$ is a question which needs to be investigated by the social scientist.[13]

Perhaps enough has been said at this point to warrant dropping c) as a sufficient condition for membership in $\underline{H}$ or as sufficient condition for inclusion in $\underline{h}$. All utterances such as the $\underline{h}^*_i$'s which contain neither replicas of nor expressions logically equivalent to the initial utterance $\underline{h}$ are to be considered as values of the variables '$\underline{x}_1$' ... '$\underline{x}_n$', and not as parts of $\underline{h}$ nor as members of $\underline{H}$. Therefore, let

'$\underline{h}$'	stand for	a particular utterance or an individual composed of a number of utterances containing $\underline{n}$ members (where '$\underline{n}$' is any integer) each of which contains as a sentential part a) a replica of the sentential parts of each of the others, or b) that which is logically equivalent to the sentential parts of each of the others. And further, that each such utterance constitutes a forediction that $\underline{e}$.

The translation criterion is dropped, and translations of individuals composing $\underline{h}$ are construed as values of one or another '$\underline{x}_i$'.

---

[13] Some of the problems confronting the social scientist in attempting to answer such questions will be considered in Chapter V.

One further word needs to be said concerning the values of an '$\underline{x}_i$'. We have remarked that any translation of $\underline{h}$ which is not logically equivalent to $\underline{h}$ is to be treated as a value of some '$\underline{x}_i$' according to the requirements of a scientific explanation outlined in Chapter I. Thus a value of an '$\underline{x}_i$' may be, but it need not be a linguistic event. It may, for example, be a certain other psychological event which is causally connected with particular utterances. These psychological events may, in turn, be "the cause" of some other utterances. What is to be observed is that the values of '$\underline{h}$', '~$\underline{e}$' and '$\underline{x}_i$' are spatio-temporal events. The substituends of these variables will be the designators of these events.

A rather peculiar relationship exists between $\underline{h}$ and ~$\underline{e}$ which demands some clarification. '~$\underline{e}$' takes as substituend a replica of the negation of the sentential part of the value of '$\underline{h}$'. There may appear to be a hint of a violation of type theory here. Events and names of events both appear on the same type level. That this difficulty is only apparent may be made clear by way of an example.

Suppose an economist $\underline{A}$ predicts at time $\underline{t}_o$ that a particular stock $\underline{S}$ will advance one point between time

t and t_n.[14] And suppose that as a causal consequence of this prediction, the stock fails to advance.[15] The substituend of 'h' will be the expression:

> The utterance by A, at t_o, 'S will advance one point between time t and t_n'.

The substituend of '~e' will be the expression:

> S will not advance one point between time t and t_n.

The fact that there are certain physical similarities between the substituends of 'h' and '~e' does not create any type difficulties. It is true that the quoted expression in h appears in the meta-language and the expression which described ~e appears in the related object language. But both h and ~e appear on the same linguistic level. The quoted expression embedded in h can actually be eliminated. 'h' could take as substituend the expression:

> The utterance of A at t_o.

This effectively denotes the utterance although it does not show what it is. The quoted expression which appears in 'h' serves to characterize the relationship which the utterance has to the event ~e. It is inserted in h

---

[14] The verb 'will advance' is here understood as tenseless.

[15] The example, but not the analysis, is due to Popper.

simply for convenience; i.e., to make clear that one of the conditions for h being self-vitiating is fulfilled. But even this point can be made clear by other circumlocutions. 'h' could take as substituend:

> The utterance of A at t_o which predicts e.

Or 'h' could take as substituend:

> The utterance of A at t_o which predicts that S will advance one point between t and t_n.

The convention of indirect discourse here effectively removes any hint of a violation of language levels. Thus the kinds of contradictions which grow out of such linguistic entities as the Grelling paradox, do not arise in our definition of 'self-vitiating'. What has been construed as a logical difficulty in regard to the phenomena of self-vitiating hypotheses does not, at least, depend upon this particular feature of the phenomena in question.

The convention adopted for the predicate 'logically implies' is for the purposes of keeping the various language levels clear. Instead of the symbolism 'Iyh' we could have used the more familiar 'y$\rightarrow$h'. Our convention, however, makes obvious the relationship between the two-place predicate 'I' and the one-place predicate 'F' (which stands for 'is false'), and the predicate 'V'. It is hoped that our convention makes this fact clear.[16]

---

[16] We write 'Fh' for the expression 'h is false'. We could not write '~h' for this would presumably have

Some interesting and important consequences of our analysis follow. First of all, it is only specific utterances which are self-vitiating. For any H only particular members of H can truly have the predicate applied to them. (In the literature, this point has not, to my knowledge, been clearly made.) 'Self-vitiating', as so far construed, is a predicate which applies to objects or events which also happen to be linguistic entities or events. We note here an important difference between 'self-vitiating' and such predicates as 'logically implies'

---

meant something entirely different from what we intend. It will be remembered that 'h' takes as substituends the names of particular expressions. Thus '~h' would either be not well-formed or else would refer to the logical complement of h neither of which we intend. Following Tarski, we may write

> 'Snow is black' is false if and only if snow is not black.

Now h(sp) will be expressions similar to those of "'Snow is black'" rather than 'Snow is black'. Hence if we wrote '~"Snow is black"' this expression would either refer to the logical complement of that which is named by the name '"Snow is black"' or else it would not be a well-formed expression. It is for this reason that we have chosen the convention of applying the predicate 'is false' to an utterance and writing 'Fh' for the statement 'h is false'.

When we wish to contrast a pair of utterances containing sentential parts which are mutually contradictory we shall write 'h' and '$\tilde{h}$'. Here what we intend is that the sentential part of $\tilde{h}$ is the negation of the sentential part of h. The reasons for this convention are similar to those just stated.

On the other hand, when we write '$\tilde{e}$', this expression is to mean the same as '~e' and is used only for convenience where 'e' is to appear as a constituent of a more complex expression.

or such predicates as 'is true' or 'is false'. Given any pair of tokens h(sp) and l(sp), if l(sp) logically implies h(sp) then for any other pair of tokens h'(sp) and l'(sp) if h'(sp) is a replica of h(sp) and l'(sp) is a replica of l(sp) then l'(sp) logically implies h'(sp). But if a particular h is self-vitiating it by no means follows that every member of H is self-vitiating. We can, of course, use some locution as 'Every member of H is derivatively self-vitiating provided that some member of H is self-vitiating.' This would make clear that we were not applying the predicate 'self-vitiating' to utterances other than those which begin the causal sequence in question. But, I believe, there are good reasons for avoiding this locution. What we wish to make clear is the distinction between those hypotheses which are simply false, and those which are false due to the fact that they enter into a particular causal relationship which brings about a state of affairs which exhibits their falsity. This distinction, we maintain, can be upheld if we distinguish between those members of H which are part of the causal sequence and those which are not (e.g., members of H uttered after ~e).

Another consequence of our analysis shows that we can assert unconditionally:

$$(h)(Vh \supset (x)(x \epsilon H \supset Fx))$$

That is to say, if any member of H is self-vitiating, then every member of H is false.

Another point has to do with "replicahood" of utterances as a condition for membership in a particular $\underline{H}$. Consider a pair of utterances $\underline{h}_i$ and $\underline{h}_j$ each of which contains a sentential part which is a replica of the sentential part of the other. This is not a sufficient condition for their joint membership in $\underline{H}_k$. One of the requirements for such joint membership would be that both $\underline{h}_i$ and $\underline{h}_j$ contain as sentential parts expressions whose negations describe $\sim\underline{e}$. The following example may elucidate this point:

> John Doe: 'I will eat the last ice-cream bar in the freezer.'
>
> Mary Doe, sister of John Doe, overhears her brother, runs to the freezer, saying, 'I will eat the last ice-cream bar in the freezer.' And she promptly eats the ice-cream.

Although the two "predictions" are replicas of each other they are not members of the same $\underline{H}$. The indicator term 'I' in the first sentence refers to John, and the second occurrence of the indicator term refers to Mary. Here we may note that, in some presystematic sense, the first utterance is self-vitiating while the second is not. Similar remarks could be made concerning elliptical expressions using tensed verbs where the time of the utterance and the time of the event is not clearly specified.

Summary

In this chapter we have attempted to show how the term 'self-vitiating' may be clearly explicated. We have

shown how certain recalcitrant utterances (such as h* and l*) may be handled without doing violence to our usage of terms such as 'logical implication', 'scientific prediction' and the like. Before proceeding to an analysis of the associated disposition predicate 'self-vitiatable' it will prove helpful to consider several related problems, among which will be the analysis of the notion of a self-fulfilling hypothesis, an examination of certain problems connected with self-fulfilling and self-vitiating generalizations, and an attempt to show how the phenomena of self-vitiation and self-fulfillment may be subsumed under a class of phenomena which have to do with unexpressed or tacitly assumed auxiliary assumptions.

CHAPTER IV

SELF-VITIATING AND SELF-FULFILLING HYPOTHESES COMPARED

Introduction

We need now to examine a phenomenon which is closely related to that of self-vitiating utterances and which presumably would need to be accounted for in an adequate empirical theory of self-vitiation. We shall attempt to determine whether the mode of analysis used in explicating the term 'self-vitiating' will also be adequate for explicating the term which will characterize the phenomenon of self-fulfilling utterances. The logical relationships between the two predicates will be examined. And finally we shall attempt to show that the phenomena of self-vitiation may be treated as subsumed under a more general class of fore-dictions whose failure is due to the lack of certain statements in the explanans or to the lack of certain qualifications of statements in the explanans.

A Definition of 'Self-fulfilling'

Another kind of hypothesis which is closely related to those of the self-vitiating type is variously called 'self-justifying belief', 'self-justifying prediction', 'self-fulfilling prediction', etc. We propose the term 'self-fulfilling hypothesis' to distinguish this kind of entity from those which we have labeled 'self-vitiating'. A few examples of this kind of hypothesis, which are in some respects analogous to those of the self-vitiating type already considered, are as follows:

A. An economist predicts that a particular stock S will increase n points or more between time t and t_0. Publication of the fore-diction causes increased demand for the stock and subsequently the stock increases n points. Circumstances are such that we may feel warranted in asserting the counterfactual: 'If the fore-diction had not been published, the stock would not have increased n points'.

B. A student B is discouraged due to his previous failure to master material M. A teacher T "predicts" in the presence of B that B will master M by time t. Having his self-confidence restored, B then proceeds to master M by time t. Again we may feel warranted in asserting a counterfactual in regard to this utterance: "If T had not 'predicted' B's mastery of M, B would not have mastered M by time t".

C. A sociologist predicts that a small group of persons G will exhibit behavior B under circumstances C at time t. The individual members of the group being in a cooperative mood proceed, as a group, to exhibit behavior B when circumstances C obtain at time t. This type of phenomena is often referred to in popular parlance as "the power of suggestion".

D. A number of interesting examples related to C. are to be found in the practice of hypnosis. A hypnotist, X, "predicts" of a particular subject, S, a heavy smoker,

that S will not smoke for a period of time t. The "fore-diction" is made in the form of what is called 'a post-hypnotic suggestion' and is made while S is in a hypnotic sleep. On being awakened S resists all attempts to get him to smoke during t.[1] In this example we may have ample warrant for claiming that if the "fore-diction" had not been uttered by X, S would have smoked during t.

Several important differences appear between self-vitiating and self-fulfilling hypotheses. But the method for analysing the manifest predicate 'self-vitiating' turns out to be adequate for the analysis of the manifest predicate 'self-fulfilling'. Let the symbols employed for the definition of 'self-vitiating' be used in the definition of 'self-fulfilling'; and in addition let:

'SFh'	stand for	'h is self-fulfilling'.
'Th'	stand for	'h is true'.
'e'	stand for	The event such that the substituend of 'e' is a replica of h(sp).

Then the predicate 'self-fulfilling' may be defined contextually as indicated by the following schema:

$$\begin{aligned} SFh =_{df} (\exists x_1)(\exists x_2) \ldots (\exists x_n)(\exists H)(\exists L)(h \in H \,.\, Chx_1 \,.\, \\ Cx_1x_2 \,.\, \ldots \,.\, Cx_ne \,.\, Th \,.\, ((y)(y \in L \supset Iyh) \\ \vee (y)(y \in L \supset Iy\tilde{h}))) \end{aligned}$$

---

[1] We shall not attempt to enter any of the intricate problems involved in distinguishing conscious from unconscious motivation. But we may suggest that a more adequate

In English this may be read: "To call a particular utterance $\underline{h}$ self-fulfilling is to say that $\underline{h}$ is part of a causal sequence which eventuates in $\underline{e}$; that $\underline{h}$ is true; and that either $\underline{h}$ or an utterance $\underline{\tilde{h}}$ is deducible from a set of utterances $\underline{L}$.[2]

The formal character of this definition differs from that of the definition of 'self-vitiating' in regard to the disjunctive expression in the definiens. This expression asserts that either $\underline{h}$ or $\underline{\tilde{h}}$ is deducible from a body of scientific laws.[3] This element of the definiens enables us to speak of an hypothesis as self-fulfilling when either a) it is deducible from a set of law-statements and antecedent-condition-statements and enters causally into a sequence of events which culminates in the event described by the hypothesis; or b) its contradictory is deducible from a set of law-statements and antecedent-condition-statements and <u>it itself</u> enters into a causal sequence which culminates in the event described by the hypothesis.

It may appear, at first glance, that this definition includes within it an assertion which is logically equivalent to the definiens of the definition of 'self-vitiating'. But this seeming equivalence is deceptive. The difference

---

understanding of the character of hypnotic suggestion would be of considerable help in the development of a theory of self-vitiation.

[2]For the usage of the negation sign above the '$\underline{h}$' see foot-note 16, p. 90.

[3]That this expression is not a tautology will be made clear presently.

lies in the non-logical rather than in the logical characteristics of a particular utterance. Suppose, for example, that according to our definition of 'self-vitiating' we have grounds for applying this predicate to a particular utterance $\underline{h}$. Are we then able to assert that, since $\underline{h}$ is self-vitiating, $\underline{\tilde{h}}$ is self-fulfilling? We are not. And this is the case because even though $\underline{h}$ may have entered a causal sequence in such a manner as to bring about the event $\sim\underline{e}$, this is not sufficient grounds for the assertion that $\underline{\tilde{h}}$ has entered any such causal sequence. If $\underline{\tilde{h}}$ has proved to be self-fulfilling, this fact will be described by an empirical assertion which is neither logically equivalent to nor deducible from the quite different assertion that $\underline{h}$ is self-vitiating.

A second point needs to be noted. The disjunct in the definiens is not a tautology. It does not, as does a tautology, exhaust all the cases. It does not, for example, take into consideration those cases such as the $\underline{h}^*_i$'s dealt with in the preceding chapter. What the disjunct does is to limit the class of utterances to which the term 'self-fulfilling' is applicable to those containing sentential parts that are well-formed. That the disjunctive expression is not required in the definition of 'self-vitiating' is seen by the following considerations: consider an utterance $\underline{h}$ such that $\underline{\tilde{h}}$ is deducible from a set of law-statements and antecedent-condition-statements; yet $\underline{h}$ is a member of a causal sequence eventuating in $\sim\underline{e}$.

In some presystematic sense we may wish to say that $\underline{h}$ is self-vitiating. Yet an important difference obtains between this $\underline{h}$ and those which conform to the requirements of the definiens of the definition of 'self-vitiating'. In regard to this $\underline{h}$ we are prepared prior to its entering the causal sequence to brand it as a false utterance. It is an uninteresting case inasmuch as it in no way adversely affects any body of theory. We shall be prepared, presumably, to hold almost precisely the same beliefs after $\underline{h}$ has entered the causal sequence as before.

But these arguments do not seem cogent in regard to self-fulfilling hypotheses. The first member of the disjunct poses problems of some importance for a theory of confirmation. When an utterance tends to bring about an event which can be predicted from a particular body of theory, to what extent can this event be counted as confirmation of the theory in question? But we need not argue this point further. If eventually it should prove feasible to include $\underline{h}_i$'s such as those characterized above within the extension of the term 'self-vitiating' we now have a means for accomplishing this. Instead of the expression '$(y)(y \epsilon L \supset Iyh)$' we include in the definiens of the definition of 'self-vitiating' the expression '$(y)(y \epsilon L \supset Iyh) \vee (y)(y \epsilon L \supset Iy\tilde{h})$'. Having taken note of this problem, we may proceed to considerations more germane to our inquiry.

'Self-vitiating' and 'Self-fulfilling' as Applicable to Generalizations

Up to this point we have tacitly assumed that self-vitiating and self-fulfilling hypotheses were singular statements. Further analysis, however, will show how utterances of generalizations may also be self-vitiating, and with an important qualification, self-fulfilling. This assumption will in no way affect the way in which these two terms have been defined. What is required is that we characterize the 'h' in our definition in such a manner as to clearly exhibit the relationships which obtain between an utterance of a universal generalization h and an event ~e. In the case where h is a universal generalization it does not appear to be an appropriate way of speaking to say that h describes a specific event, e. Instantiations of h may, however, describe e, when appropriate antecedent conditions are specified. e may, for example, be described by the statement 'Fa.Ga'. Thus if h is an utterance of the generalization '(x)(Fx ⊃ Gx)' and we are prepared to assert that Fa, then we can deduce 'Fa.Ga' from the conjunction of '(x)(Fx ⊃ Gx)' and 'Fa'.

With this example in mind we may now characterize h in such a way as to include within its scope universal generalizations. We now let:

'h'	stand for	a particular utterance or an individual composed of a number of utterances containing n members (where 'n' is any positive integer) each of which contains

> as a sentential part a) a replica of the sentential parts of each of the others, or b) that which is logically equivalent to the sentential parts of each of the others. And further, that each such utterance constitutes a fore-diction that $\underline{e}$; or $\underline{h}$ is a universal generalization such that there exists an $\underline{h}_1$ and $\underline{h}_2$, such that the conjunction of $\underline{h}$ and $\underline{h}_1$ implies $\underline{h}_2$, and either $\underline{h}_2$ or $\underline{h}_1 \cdot \underline{h}_2$ constitute a fore-diction that $\underline{e}$, it is not the case that $\underline{h}_2$ is deducible from $\underline{h}_1$ alone, and finally, $\underline{h}_1$, is true.

In a number of instances $\underline{h}_1$ will simply be an instantiation of the antecedent of the universal generalization. This need not be the case, however. And the only logical requirement is that the utterance $\underline{h}_2$ (or the conjunction $\underline{h}_1$ and $\underline{h}_2$) which describes $\underline{e}$ must be logically deducible from the generalization $\underline{h}$ when $\underline{h}$ is taken in conjunction with $\underline{h}_1$.[4]

We may now examine some of the characteristic features of the two predicates 'self-vitiating' and 'self-fulfilling'. Suppose that $\underline{H}$ is a universal generalization, one of whose utterances $\underline{h}$ conforms to the definition of 'self-fulfilling hypothesis'. That is, there is some $\underline{h}_2$ describing $\underline{e}$, which is logically deducibe from the

---

[4]For the requirements concerning instantiation which shall be tacitly assumed throughout this essay, see, Nelson Goodman, Fact, Fiction and Forecast, p. 68ff.

conjunction of $\underline{h}$ and $\underline{h}_1$. By definition, to say that $\underline{h}$ is self-fulfilling is simply to say that there is some utterance of $\underline{H}$ which causally culminates in a confirming instance of $\underline{H}$. It is not to say that the utterance of $\underline{h}$ causes $\underline{h}$ to become true. Now there may be some cases in which the locution 'the utterance of $\underline{h}$ causes $\underline{h}$ to become true' would be used without difficulty. These cases, however, would involve non-law-like rather than law-like generalizations. For example, I could spray the rug in my study with gasoline in such a manner that when I touched a match to the rug the flames would spell out a replica of the sentence 'All the fibres in the rug in A. J. Stenner's study will burn up at time $\underline{t}$'. It might be held that this is a generalization, although a non-law-like one, which on its flaming occurrence causes itself to become true in some sense of "cause". But there are sufficient reasons for avoiding such locutions. And it can be shown that other ways of speaking enable us to say what we wish to say without encountering the difficulties which arise when such locutions are employed.

Now the interesting cases we wish to consider are those law-like generalizations which are not restricted in such a way as are so-called non-law-like generalizations. And granted that establishing of a criterion for distinguishing between law-like and non-law-like generalizations is difficult, the difficulty, again, is not one which is peculiar to our subject matter. We may therefore simply

note the difficulty; for it does not seem to be one whose solution would affect our analysis one way or the other. On the contrary, it would appear that any criteria which are finally adopted for distinguishing between law-like and non-law-like generalizations would need to take cognizance of some of the points made in our analysis of self-fulfilling and self-vitiating hypotheses.[5]

---

[5] It may turn out that the kind or the amount of evidence needed to determine whether or not a given utterance of a generalization is self-vitiating or self-fulfilling will also enable us to determine whether the generalization in question is law-like or non-law-like. What is involved is the following: in the case of some generalizations the expression 'Th' in the definiens is redundant; whereas in other generalizations it is not redundant. (The proof of this assertion will be indicated presently.) That is to say, if from the other components of the definiens we are able to determine the truth of the assertion 'Th', then 'Th' is redundant and the generalization in question is one we would consider as non-law-like. If, however, 'Th' is not an expression whose truth can be determined by reference to the truth of the other components alone, then 'Th' is not redundant and the generalization is a law-like one.

Before this problem can be adequately dealt with, what is required is a definition of 'self-fulfillable' and 'self-vitiatable' both of which are disposition predicates. The reasons for our inability to define these terms at present will be given in Chapter V. But in the event that such definitions are forthcoming we may now indicate how they may help to solve the problem of determining the law-likeness of certain generalizations. Suppose that an utterance of a generalization $\underline{h}_g$ is one such that under certain circumstances $\underline{C}$ we would apply the predicate 'self-fulfilling' to $\underline{h}_g$. This is to say that $\underline{h}_g$ is self-fulfillable. Now the question as to whether or not $\underline{h}_g$ is law-like or non-law-like could be determined by deciding whether or not the expression 'Th_g' would or would not be redundant under circumstances $\underline{C}$.

The question of the redundancy of 'Th' will be in part dependent on the kind and specificity of the value of '$\underline{e}$', and the logical relationships which obtain between '$\underline{e}$' and $\underline{h}$. Consider the assertion: '$\text{Th} \equiv \underline{e}$'. In regard to

In regard to law-like generalizations the predicates 'self-fulfilling' and 'self-vitiating' are logically incompatible. It might seem presystematically that we could have two utterances of a single H one of which would be self-vitiating and the other self-fulfilling. It will indeed be the case that two utterances containing sentential parts which are replicas of each other will, or at least probably will, have different causal consequences. But we can show that this creates no particular difficulty for our analysis which is not also encountered in other theories.

Logically considered, no amount of evidence can ever completely verify a contingent universal generalization. But one disconfirming instance, if it is confirmed to a high degree, may lead us to reject the generalization, provided that "other things are equal". This is not to say that we do not have means for preserving those generalizations which have proved exceedingly useful and which we are loathe to give up as false. It is to make a simple logical point: namely, that the assertion

---

a generalization h_g if we are prepared to assert that h_g is true if and only if è then we will also be prepared to assert that h_g is non-law-like. For presumably law-like generalizations do not "describe" e in the same manner that non-law-like generalizations do. The problem then boils down to two questions: 1) What can serve as the value of 'e'? and 2) What constitutes confirmation of e? We can answer the first question; only actual events may serve as values of 'e'. The answer to the second question waits upon an adequate theory of confirmation.

'Fa . Ga' is logically incompatible with the assertion '(x)(Fx ⊃ ~Gx)'. And if we wish to hold that the former is true we must then hold that the latter is false.[6] Whether we give up the generalization or find some means of modifying it, or whether we discount the evidence for the singular statement will depend on considerations which are other than logical ones. The point we wish to make is logical. According to the way in which 'self-vitiating' and 'self-fulfilling' have been defined, a particular hypothesis h cannot have both predicates truly applied to it. And it follows that there cannot be two members of a particular H such that one is self-vitiating and the other self-fulfilling. If an hypothesis is self-vitiating, then by definition it is false; and we noted that if one member of H is false, then every member is false. On the other hand, if a particular hypothesis is self-fulfilling, then by definition it is true. It follows, therefore, that there can be no two members of an hypothesis one of which is self-fulfilling and the other self-vitiating.

From this analysis an important difference between the two predicates emerges. The grounds on which a particular utterance of a law-like generalization can have the predicate 'self-vitiating' applied to it are

---

[6]Unless, of course, we are so bold as to be willing to give up the law of contradiction.

clear, from a logical standpoint: one disconfirming instance of the generalization which conforms to the requirements of the definiens of the definition will be sufficient to apply the predicate 'self-vitiating' to the hypothesis. But one confirming instance of a law-like generalization will not be sufficient evidence for applying the predicate 'self-fulfilling' to a law-like generalization.[7] We need to know on other grounds that the hypothesis is true; for just as one swallow does not make a summer, neither does one confirming instance of a universal generalization render that generalization true. One of the consequences of this fact is that to say that a particular utterance of a generalization is self-fulfilling, is to say that it never has had, does not now have, and never will have either a disconfirming instance or, more narrowly considered, a self-vitiating instance. But to call an hypothesis 'self-vitiating' is not to

---

[7]We are, of course, overlooking some notorious difficulties in confirmation theory. Our willingness to accept an assertion such as 'Vh' will depend on a great many factors among which will be the strength of the evidence for each of the assertions in the definiens and the cost of abandoning h as a useful generalization. These difficulties do not appear to be ones which will adversely affect our analysis. We are concerned specifically with the problem of what follows from the truth of assertions such as 'Vh' or 'SFh'; and not the problem of what would constitute confirming evidence of these assertions. The latter problem is one which would require the cooperative endeavors of a number of the sciences and a rather highly developed theory of self-vitiation as well as an adequate theory of confirmation. We may perhaps then be pardoned for apparently ignoring these issues.

claim that the hypothesis has not had any confirming instances. On the contrary; presumably the overwhelming majority of generalizations to which we would be willing to apply the predicate 'self-vitiating' will be generalizations which have been confirmed to some degree and in many cases have been confirmed to a high degree.

To see what is involved we may consider an example frequently used in the literature:

E. All ravens are black.

We may inquire as to some of the conditions under which this generalization may be self-vitiating or self-fulfilling. There are several ways, in fact a large number of ways, in which the generalization could be self-vitiating. An utterance of E. could be inscribed on the body of a female raven R. The instruments used for the inscribing could be tatooing needles containing some radioactive substance S. S we assume is of such a nature that its presence in the body of ravens brings about certain genetic changes in the offspring of R, i.e, the offspring of R are not black. Hence the utterance of E. would be self-vitiating.

Another way in which E. could be self-vitiating would be in the event that tokens of E. were produced by spraying a chemical S' into the air in the vicinity of R's nesting place. The chemical S' would be one which changed the nature of the pigmentation in the feathers

of R and the causal consequence of this spraying would be that R was no longer black. A more fertile imagination than our own could undoubtedly conceive of many other possibilities.[8]

We may also inquire concerning the conditions under which utterances of E. may be self-fulfilling. (We need to keep in mind that one of the requirements of a self-fulfilling hypothesis is that it be true.) An interesting example, and one which also shows some of the relationships which obtain between the predicate 'self-fulfilling' and the predicate 'self-vitiating' would be the following: raven, R, previously innoculated with chemical, S, is innoculated with a new substance, S'', which counteracts the "possible" effects of S prior to the actual birth of any offspring of R. The innoculation, as in the previous instance, is done in such a way as to spell out tokens of E. The causal consequences of these innoculations of R with S'' is that all of the offspring of R are born

---

[8] If these examples seem remarkably far from what we would usually construe as self-vitiating hypotheses we may indicate some our reasons for choosing these examples. First, they are simple enough to avoid the kinds of complications which arise in regard to purposive behavior. Second, they serve to show that the use of the phenomena of self-vitiation as a means for distinguishing the social from the physical sciences is without forundation, inasmuch as our "far-fetched" examples conform to both our pre-systematic discussion of self-vitiating and our systematic requirements. Third, the discussion based upon these examples serves as an introduction to the problems discussed later in the chapter under the heading of 'Boundary conditions'.

black and remain black. In this latter instance the second utterance of E. interrupts the causal sequence which was begun by the first utterance of E. Hence, provided that E is true, the predicate 'self-fulfilling' may be applied to this second utterance of E. In such an event the predicate 'self-vitiating' would not be truly applicable to the first utterance of E.

The underlined expression in the above paragraph describes a condition which is of crucial importance in distinguishing the predicates 'self-vitiating' and 'self-fulfilling' as applied to generalizations. It may be shown that the expression 'Fh' in the definiens of our definition of 'self-vitiating' is redundant. That is, if we are able to assert that the other expressions in the definiens are true, then we can show that 'Fh' logically follows from these other assertions and that hence we may apply the predicate 'self-vitiating' to h.[9]

---

[9]The proof is as follows:

Let 'Chẽ' stand for '(x)(Gx ⊃ Kx) . (x)(Kx ⊃ ~e) . Gh'

1.	'~e' is true ≡ ~e	By semantic definition of 'true' and 'false'.
2.	h_2 is false ≡ '~e' is true	By semantic definition of 'true' and 'false' and the relation of h to ~e.
3.	Chẽ	Premise
4.	(x)(Gx ⊃ Kx) . (x)(Kx ⊃ ~e) . Gh	By replacement for 3.
5.	Gh ⊃ Kh	4, simp., spec.

The reason for keeping 'Fh' in the definiens despite its redundancy is that of preserving a notational symmetry between the predicate 'self-fulfilling' and the predicate 'self-vitiating'. But the expression 'Th' in the definiens of the definition of the predicate 'self-fulfilling' is by no means always redundant. That this is the case is readily seen by observing that being able to assert 'Fa.Ga' is not sufficient grounds for asserting '(x)(Fx ⊃ Gx)'.

The distinction we are here considering between the two predicates 'self-vitiating' and 'self-fulfilling' may be summed up in the following way:

1) If the assertion '$(\exists x_1)\ldots(\exists x_n)(\exists L)(\exists H)(h \epsilon H \,.\, (y)(y \epsilon L \supset Iyh) \,.\, Chx_1 \,.\, \ldots \,.\, Cx_n\tilde{e})$' is true, then 'Fh' is true. And hence 'Vh' is true.

2) If the assertion '$(\exists x_1) \,.\, \ldots \,.\, (\exists x_2)(\exists L)(\exists H)(h \epsilon H \,.\, Ch_{\sim}x_1 \ldots \,.\, Cx_n e \,.\, ((y)(y \epsilon L \supset Iyh) \vee (y)(y \epsilon L \supset Iyh)))$' is true, it may or may not be the case

---

6.	$Kh \supset {\sim}e$	4, simp., spec.
7.	$Gh \supset {\sim}e$	5, 6, syl.
8.	Gh	4, simp.
9.	${\sim}e$	7, 8 m.p.
10.	'${\sim}e$' is true	1, 9, subst.
11.	h_2 is false	2, 10, subst.
12.	$(h \,.\, h_1) \rightarrow h_2$	Premise
13.	$(h \,.\, h_1)$ is false	11, 12, by definition of '$\rightarrow$'
14.	h_1 is true	Premise
15.	h is false	13, 14, by definition of '.'

that 'Th' is true and hence it may or may not be the case the 'SFh' is true.

It can be seen that different grounds are therefore necessary for determining whether the two predicates 'self-fulfilling' and 'self-vitiating' will apply in any given case of a generalization.

Up to this point we have been considering 'SF' as applied to generalizations. When we come to the consideration of singular statements, however, we discover that the difference between the two predicates just mentioned does not obtain. For in the case of an utterance of a singular statement we discover that the expression 'Th' in the definiens of the definition of 'self-fulfilling' is also redundant. The proof for this is similar to the one given for 'self-vitiating'. We need simply note that if the event predicted by an utterance of h comes about, then by the semantic definition of truth we are able to assert that h is true.

We may now inquire as to whether or not there are any self-fulfilling generalizations such that the expression 'Th' in the definiens is redundant. The answer of course will be dependent on a vast array of empirical evidence, evidence which presumably we do not have at our disposal at present. But we can easily see how 'Th' may be redundant in a given instance by considering an example previously dealt with in regard to the burning of all the fibres in a particular rug. In this example the value of

the sentential part of h is a replica of the substituend of 'e'. When this particular relationship obtains between h and e, in this instance we are able, by the semantic definition of truth, to deduce 'Th' from the other components in the definiens.[10]

This example differs in an important respect from another which shares certain features with it. Consider a case in which raven R is innoculated with a substance S*. The result of this innoculation is that a mutated species is produced which we shall call 'croven'. The innoculation is again performed with tatooing needles spelling out an inscription which is a replica of:

F. All crovens are blue.

We assume further that there are good empirical grounds for believing that all crovens are in fact blue. Now suppose that after three generations of crovens all the crovens thus born of R are destroyed. Are we in a position to claim that the assertion "'All crovens are blue' is true" is redundant in the assertion that F is self-fulfilling? The answer to this question appears to be in the negative.[11]

---

[10] See footnote 5, pp. 104-105 of this essay for the relevance of this distinction to the problem of distinguishing law-like from non-law-like generalizations.

[11] Unless, of course, we wish to define 'croven' as 'decendent of R' rather than specifying other characteristics as the defining characteristics of croven. But such a move does not appear to be one which has proved the most fruitful in the sciences.

One reason for this fact would be that the example does not rule out the possibility that other crovens may be born under similar circumstances or the offspring of another raven R' may mutate by "natural" means. In the case of the rug fibre example then we are prepared to claim that the 'Th' in the definiens is redundant, whereas in the croven example we are not prepared to make this claim. Wherein lies the important difference?

As we have previously noted, in the rug fibre example the value of h(sp) is a replica of the substituend of 'e'. But in the croven example this relationship does not obtain. Speaking loosely, we suggest that in the rug fibre example the generalization describes only actual fibres whereas in the croven example the generalization speaks "about" possible as well as actual crovens.

We are not, of course, specifically concerned to solve the problems encountered in a characterization of law-like generalizations. We have merely been attempting to note that this problem is not unrelated to the problems encountered in an analysis of self-vitiation. And presumably the solution of the one will contribute to the solution of the other. In the next section we shall attempt to show how the problems of self-vitiation are directly related to another, broader class of problems in the sciences.

"Boundary Conditions" and Self-vitiating Hypotheses

In this section we wish to show that the class of self-vitiating and self-fulfilling hypotheses can be construed as a sub-class of a certain type of phenomena. The characteristic feature of this class of phenomena concerns the nature of the so-called boundary conditions which must be fulfilled if the event is to be successfully predicted from a body of law. The hypothesis to be advanced in the course of our argument is that the same characteristic features are present in those utterances which we have called 'self-vitiating' and 'self-fulfilling' as are present in other predictions whose success or failure is dependent on making the boundary conditions explicit.

In our example of the disruption of the orbit of Mars[12] we noted how the production of tokens of a certain type could bring about an event such that the fore-diction exhibited by the tokens turns out to be false. Now it can be easily seen that in the physical sciences, at least (and we hope to show that the same will hold true for the social sciences) this class of phenomena is easily subsumed under a more general class of phenomena. What occurs in these cases is that the so-called "boundary conditions" are not clearly specified in the theory in question but are tacitly assumed in any scientific explanation of

---

[12]See pp. 34-35.

events or in the deduction of certain general laws from more general laws. A few examples may serve to show how, in the physical sciences, instances of self-vitiating hypotheses may be subsumed under this more general class.

Consider the following deductive explanation for the fact that a particular iceberg, B, floats:

G. All ice will float in water.
All icebergs are ice.
B is an iceberg.
B is in water.

∴ B floats

Now suppose that there is a particular object O, which we would designate as a glacier, which has, in moving down a mountain side in Alaska, picked up a large amount of glacial debris which is embedded in its interior. O slides into the Pacific Ocean, and as it moves into the Japanese currents begins to melt. As it melts it sinks to the bottom of the ocean due to the large quantities of glacial debris embedded in its interior and the consequent increase in the mean density of O as it melts.

From the set of premises in G., together with the premise 'O is an iceberg' we can deduce that O floats when placed in water. Now this conclusion contradicts what by assumption are the facts. Hence we must modify or reject at least one of the premises. There are a number of ways in which the required modifications can save the laws stated in the premises. One such way would be to define 'ice' so that an object such as O would fail to qualify

as ice. For example, we could specify that an object would be considered to be ice only if its mean density was less than that of water. O would then not be treated as an instance of ice. It also could not be considered as an instance of an iceberg. This might place seemingly artificial restrictions on popular ways of speaking, but such speech, as is indicated by the problems arising from example G., often need clarification if we are to avoid the kinds of puzzles which thwart our attempts to understand and explain phenomena.

Another method of saving the appearances would be to deduce the law 'For all x, if x is ice, then x has a mean density less than that of water' from physical theory. O may then be excluded as being an instance of ice on the grounds of its failure to conform to the law in question. That is, by modus tollens we conclude that O is not ice. But this method possibly can lead to contradictory results when taken with other law-like statements of physics and chemistry. Consider

J. All water freezes (i.e. becomes ice) at 32°F or below.
O is composed of water and its temperature is 29°F.
29°F is below 32°F.

∴ O is ice.

Since the conclusion to example J. contradicts our former conclusion that O is not ice, other modifications need to be made in the theory. We can, for example, define

'water' in such a way that a particular sample of liquid will be called 'water' only if the sample in question contains no more than a specified proportion of a foreign substance, i.e. a substance other than one whose chemical properties are characterized by the formula 'H_2O'. Another way of proceeding, one probably more in accord with actual chemical practice, is to define water simply as 'H_2O'. The fact that there are few if any samples of this pure mixture in nature will demand that some other set of auxiliary assumptions be included in any predictions about the behavior of water under actual empirical conditions.

In many cases these auxiliary assumptions are not required for the successful predictions of the behavior of water under certain conditions. Auxiliary assumptions are usually required in cases similar to those of our example of the iceberg. The notorious ceteris paribus clause can always be invoked when recalcitrant phenomena such as the occurrence of O becomes problematic. But to invoke such a clause is seemingly only to make clear that the necessary auxiliary assumptions or definitions have not been included in the premises from which the explanandum was deduced. Once these auxiliary assumptions have been introduced, the falsity (or perhaps the ambiguity) of one of the premises is determined. And from the revised set of premises the false explanandum is no longer deducible.

With these points in mind we may now consider another example. Instead of an iceberg, let 'O'' stand

for an object which has been formed in the following manner: a number of pieces of lead buckshot are placed on the bottom of a refrigerator tray, arranged in a manner to spell out tokens which are replicas of:

K. O' will float in water.

The refrigerator tray is then carefully filled with water and placed in the freezing compartment of a refrigerator. When the liquid has frozen solid, it is removed together with the embedded buckshot and placed in a tub of water. As the mean density of the composite object O' is increased due to the melting of the ice, O' subsequently sinks to the bottom.

It can be readily seen that the generalizations functioning as auxiliary assumptions required to explain the fact that O did not float are precisely the generalizations required to explain the fact that O' did not float. One difference between O and O' concerns the arrangement of the non-aqueous material contained within the two objects. The glacial debris in O presumably is not arranged in a manner such that the debris constitutes a prediction of the behavior of O; while the lead shot in O' is arranged to spell out a prediction of the behavior of O'. But even this difference need not obtain. Statistical theory would lead us to assert that given an indefinitely long period of time and the production of a vast number of icebergs, there is an overwhelming probability that there would appear an iceberg which exhibited the characteristics of

O and in addition had its debris arranged in such a manner as to spell out a set of tokens which were replicas of the sentence 'O'' will float in water.' Let us call such an object 'O'''. The explanans which is required to explain why such objects sink when immersed in water contain the same general laws in the cases of O, O' and O''. The fact that the differences between these three objects may be important for the explanation of other characteristics of these objects should not obscure the fact that in the particular explanantia we are here considering these differences are irrelevant.

Perhaps it will be objected that in the social sciences the shape of the tokens involved constitutes one of the relevant factors in any explanation of self-vitiating hypotheses, whereas, in the examples just given, the shape of the tokens obviously is unimportant for predicting the behavior of the objects in question. For example, if the United States' State Department predicts:

L. Russia will attack West Berlin at time t.

this utterance and others containing as sentential parts replicas of it will have, presumably, vastly different causal consequences than a set of utterances of the form:

L'. West Berlin will attack Russia at time t.

Here it is the difference in the order (or what we might call 'the shape') of the tokens which will determine the actual outcome of events. No theory of self-vitiation which failed to make this point clear could possibly have predictive value. L' since it contained all and nothing but replicas of the word-tokens in L would be treated as having the same causal consequences as L which we know presystematically is not likely to be the case. Now whereas in the cases of O and O' the arrangement of the debris is an irrelevant factor in explaining the causal consequences of the ice's containing foreign objects, in the case of the utterances of L and L' the arrangement of the tokens is of crucial importance.

In replying to these objections we may agree that any theory of self-vitiation will need to make clear when the shape of the tokens is relevant and when irrelevant. But it is by no means clear presystematically what is to constitute the relevant shape characteristics of the objects in question. Suppose, for example, that the iceberg O'' contains all of the debris embedded in the outer layer rather than within the central portion. Then when the iceberg, O'', reaches the warm waters of the Japanese currents, the glacial debris will begin to drop off so that the mean density of the iceberg remains less than that of water and O'' continues to float rather than sink to the bottom. Here we can see that the shape of the debris is of considerable importance to the outcome of the predicted event.

It may be objected that what is involved here is not so much that of shape but that of the position of the debris, whereas in the case of predictions L and L' what is involved is the shape of the tokens, and that hence we need to distinguish between those predictions such as would be falsified by an appearance of an object such as O'', in which the important characteristic is the position of the debris and not the shape. Ignoring the obscurity of the objection for a moment, we may note that the relative position of utterances L and L' would also be an important consideration in any attempt to predict what the causal consequences of L or L' might be. If the only extant replica of L occurred on the head of a pin or was inscribed on the underside of President Kennedy's desk, the causal consequences of such an utterance of L would be presumably much different than if the replicas of L appeared in prominent positions on the front pages of the leading newspapers of the world.

Our example of disrupting the orbit of Mars points up also how important the shape of the tokens may be in determining whether or not a prediction of a physical event is self-vitiating. Suppose that the cobalt bombs are built to explode in such a fashion as to spell out a sentence token which is a replica of:

M. Mars' orbit will not be altered by a factor x at time t.

Suppose further that there is one bomb for each of the word-tokens and one for each of the substituends of the simple literals '_x_' and '_t_', and that these bombs must explode in a precise order such that if they explode in any other order the alteration of Mars' orbit will not take place. Now whether or not the fore-diction in question will or will not be self-vitiating will depend principally on the shape of the tokens produced by the exploding bombs.

It seems therefore clear that position, shape, and order of tokens are or may be equally important in a self-vitiating hypothesis in physics as in a self-vitiating hypothesis from sociology or psychology. These characteristic features of the sentential parts of self-vitiating utterances, therefore, are not those which enable us to make any significant discrimination between the social and the physical sciences.

A further objection might take the following form: it is granted shape and position of the tokens may be factors in predictions of events whether these have to do with the behavior of icebergs and planets or the behavior of human beings. But there is one important difference between social and physical phenomena which has been overlooked. In the case of the cobalt bombs the same effect could be achieved by means other than that of having the bombs spell out the particular tokens in question. Whereas in the case of _L_ the causal consequences are dependent on certain _meanings_

attached to the tokens, in the case of disrupting the orbit of Mars meanings are irrelevant. The bombs could explode in the shape of an American Flag, or a picture of Santa Claus, and provided that the bombs exploded according to certain laws of mechanics the result would be the same whether or not they exploded as tokens of a fore-diction or whether they exploded according to some other pattern. But in regard to the prediction concerning Russia's behavior in West Berlin the tokens must be replicas of L or similar to L, otherwise the causal consequences will not be the same as they would be if utterances of L. were alone in question.

Now we need not, I believe, resort to "meanings" in order to deal with the problem in question. The force of the objection is lost due to the obscurity of the term 'similar'. By introducing the notion of tokens which are similar to L one only begs to question as to how similar the tokens will need to be in order for them to have the same causal consequences as L. We need also in the case of Mars' orbit to consider how similar the proposed different explosions need to be to the ones initially proposed in order for the same causal consequences to result.

If one does not resort to the notion of 'similar to' and instead attempts to draw the distinction on the basis of the "meanings" and the causal consequences of these "meanings" it is by no means clear that what can

be explained by an appeal to the meanings of the tokens cannot also be explained without resorting to such entities. If what one wishes to do is to distinguish between the consequences of linguistic events involving purposive behavior and those involving icebergs, ice-cubes and exploding bombs, there may be certain distinctions which can be made. We shall consider some of these in the next chapter. But it is not prima facie obvious that an appeal to meanings helps clear up any obscurity in the distinction. As we noted before, the paucity of our imagination may be the only thing which is lacking from attributing meaning to any object or event whatsoever.[13] That is to say, any event or object may be interpreted as a sign which indicates some other event. In some construal of 'meaning' which is yet semantical (rather than valuational in the sense of 'importance') any event whatsoever can have a meaning.

The point of these remarks is <u>not</u> that the designative function of utterances has no bearing on the outcome of events generally nor on the phenomena of self-vitiating sign-events in particular. All we are

---

[13]"Whatever the author intended the token to be a clew to, that is what the token means." Henry Leonard, <u>Principles of Right Reason</u>, p. 138 Cf. Also, Lewis Carroll, <u>Through the Looking Glass</u>, (Random House, New York, 1946, special edition) "'When I use a word,' Humpty-Dumpty said, 'it means just what I choose it to mean - neither more nor less.'" Analogously we may extend the notion of 'token' so as to include objects or events of any size, shape or description whatsoever.

suggesting at this point is that the introduction of "meanings" into our universe of discourse does not help but may actually hinder our understanding of the problem of self-vitiation. This problem we take to be that of the development of a theory which will enable us to explain (in the sense explicated in Chapter I) occurrences of self-vitiating hypotheses.

Summary

In this chapter we have attempted to explicate the notion of 'self-fulfillment' in a manner directly analogous to that of the notion of 'self-vitiation'. We have found that the method of analysis used for explicating the latter is adequate to the analysis of the former. We went on to show some of the logical relationships obtaining between the two terms. And finally, we tried to show how the phenomena of self-vitiation may be understood as subsumed under a more general class of phenomena in which prediction of the future is unsuccessful due to failure of the prediction to make clear certain boundary conditions. We attempted, at the same time, to answer possible objections to our analysis.

Throughout this chapter we have been concerned primarily to explicate the manifest predicate 'self-vitiating'. Next we must consider what would be involved in a definition of the associated disposition predicate 'self-vitiatable' and related problems. It is to this task that we now turn our attention.

CHAPTER V

TOWARD A THEORY OF SELF-VITIATION

Introduction

Up to this point in the discussion we have attempted to characterize, more precisely than seems hitherto to have been achieved in the literature, the phenomena of self-vitiation. We have attempted to define a predicate which could be applied to these phenomena with a minimum risk of ambiguity and vagueness. Now we must examine a much more difficult problem, the problem of determining the conditions under which we shall apply the disposition predicate associated with the manifest predicate 'self-vitiating'. We hold that the task of adequately defining this particular disposition predicate will depend crucially on empirical considerations. But there are nevertheless certain formal problems which must be considered before it is possible to construct what would be acceptable as an adequate theory of self-vitiation. Among such considerations is the classification of various types of self-vitiation. We shall claim that for these and other purposes the distinctions which prove the most useful are not those which traditionally are held to obtain between the social and the non-social (or physical) sciences. We are primarily concerned with those types of self-vitiation in which important factors are those of goal-directed-behavior or purposive-behavior or sign-behavior.

A New Predicate, 'Self-vitiatable'

So far we have seen how the predicate 'self-vitiating' can be applied to particular utterances. This predicate is, however, applicable only in those cases in which the characteristic feature of self-vitiation is actual. That is to say, the predicate 'self-vitiating' is correctly applicable to just those utterances in which it is the case that the utterance in question is causally efficacious in bringing about ~e. We now wish to be able to describe those cases in which an actual utterance has not been self-vitiating but which nevertheless is such that we believe it would be self-vitiating if it were to be uttered under certain particular conditions. We are, in short, raising the question of how to explicate 'potentially self-vitiating' as applied to hypotheses. In raising this question we confront the problem of suitably defining an appropriate disposition predicate. The specific "potentiality" with which we shall be concerned is that particular feature of those utterances which, under certain circumstances, would enter a causal sequence such that an event described by their negations would result. The predicate which we shall use to characterize this feature of utterances is the predicate 'self-vitiatable'. In other words we wish to direct our attention to the assertion, "If there is a member h of a particular statement H such that if h were uttered under suitable circumstances, C, h would be self-vitiating, then

h is self-vitiatable." Several points need to be noted concerning this characterization of the predicate 'self-vitiatable'. We shall state each point briefly and then discuss each one in greater detail.

First, the characterization is not a formal definition. It gives one of the sufficient but none of the necessary conditions for the application of the term 'self-vitiatable'.

Second, it does not specify what are the "suitable circumstances" under which h would be self-vitiating. Presumably the specification of these circumstances would depend upon results of empirical investigation. It would, for example, be required, prior to a formal definition of the general predicate 'self-vitiatable', to examine a large number of actual utterances which were self-vitiating, to determine specific features shared in common by all of these examined instances, and then to determine which of these common features are relevant for characterizing the phenomena of self-vitiatability.

Third, the characterization, given above, is expressed in the form of a counterfactual. This characterization, of course, is not the only type available to us.[1] We could translate this counterfactual characterization

---

[1]Much of the material in this section is dependent on the analysis of dispositions given by Nelson Goodman in Fact, Fiction, and Forecast, (Cambridge, Massachusetts: Harvard University Press, 1955).

in such a way that the antecedent and consequent of the statement appear in the indicative mood. Translated in this manner, our characterization would read, "Since the statement, 'There is a member h of a particular statement H which is uttered under suitable circumstances C' is true, the statement 'h is self-vitiating' is true." This translation does not replace a counterfactual by a material conditional. The translation also has the counterfactual form for, by hypothesis, its "since-clause" is contrary to fact. But it deals with certain semantic relations which are claimed to hold between the two indicative components. "The disposition term ... is eliminated without the introduction of any such troublesome word as 'possible'; only non-dispositional predicates appear to remain, even if they are slightly jaundiced with a modal inflection."[2] This translation is one which may make us "feel that we have exchanged an ontological problem for a linguistic one".[3]

Fourth, as in the case of its associated manifest predicate, the term, 'self-vitiatable' is applicable to utterances of statements rather than to statements.[4]

---

[2] Op. cit., p. 40.

[3] Ibid.

[4] See footnote 15, p. 12 of this essay for a qualification of this assertion.

The disposition term is a predicate only of actual utterances (i.e. past, present, or future ones) which are members of H and not of possible members of H. This is a matter of some importance and it will be dealt with in more detail in a moment. We shall not ignore "the possible but non-actual" members of H but shall treat these as being in the extensions of predicates that apply to persons; i.e. as dispositions of persons. Thus, a possible, non-actual, self-vitiatable h is an actual disposition to utter self-vitiating hypotheses.

Sixth, this treatment of the predicate 'self-vitiatable' avoids any crucial ambiguity in the antecedent of the conditional characterization. Note that the counterfactual is phrased "If there is an h which is a member of H" and not "If there were an h ..."

Seventh, it follows from the fifth point that the extension of 'H' is not null. That is to say, if there is some actual utterance of h then there is at least one member of H and hence 'H' is not an empty description. We are thus able to avoid having to deal with problems created by allowing as substituends of the variables names of non-existent entities.[5]

---

[5]An example given by Henry Leonard is the following:

Santa Claus lives at the North Pole
∴ Someone lives at the North Pole

Whether one accepts this inference as valid will depend on a number of issues with which we need not be concerned. The example is from Philosophical Studies, H. S. Leonard, "The Logic of Existence", 7:52, June, 1956.

Eighth, we have been fortunate in finding a definition of 'self-vitiating' which is serviceable in the characterization of the predicate 'self-vitiatable'. This move is not always open to us in attempting to show the relationship which obtains between a specific manifest predicate and its associate disposition predicate. When we are fortunate enough to discover this relationship, we are then able to see more clearly the direction we must take in defining the disposition predicate.

Ninth, by making sign-events (rather than tokens or statements) values of our variables we are better able to deal with those situations in which there exists a relevant "time lag" between the production of tokens and the receipt of the tokens.

The Problem of Disposition Predicates

We shall assume that it is not required that we recount the history of attempts to deal with the problem of defining disposition predicates. We need only note that the material conditional and bilateral reduction sentence are as inadequate to our particular task as they are to the problem of stating the criteria for defining disposition terms generally. Since no general solution to the problem of defining disposition terms has yet been found, we have been forced to make use of the counterfactual form of statement in order to characterize the predicate 'self-vitiatable'.

But the counterfactual form itself poses some difficult problems. Several of these problems as they are related to the counterfactual we are here considering may be indicated.[6]

Part of our particular problem lies in the fact that the antecedent and consequent are both false in the counterfactual characterization of 'self-vitiatable'. But so are the antecedent and consequent of the following counterfactual:

> If h were made of blue cheese, then h would be self-vitiating.

Now in the event that we wish to accept our characterization of 'self-vitiatability' we shall have to do so on grounds other than those of the truth values of the components since the latter counterfactual also has both a false antecedent and a false consequent; and presumably we shall wish to reject the latter counterfactual as false while accepting the former as true.[7] To simplify our problem, we shall assume the existence of h and hence, the specific counterfactual we shall consider (now no

---

[6]Throughout this portion of the discussion we shall assume that we understand the nature of the "suitable circumstances, C" indicated in the characterization. Needless to say we do not know what these circumstances are. And one of the tasks for an empirical theory of self-vitiation would be to discover the nature of these circumstances. The point is that the statement, describing whatever conditions constitute suitable circumstances, need not contain dispositionals. More of this below.

[7]We are here assuming, of course, that our theory of self-vitiation will show being made of blue cheese to be unconnected with self-vitiation.

longer an "indicative" counterfactual,) is the following:

> If h were uttered under suitable circumstances, C, then h would be self-vitiating.

When we assert of any utterance h that it is self-vitiatable what we intend by this assertion is that if h were uttered under suitable circumstances, C, then h would be self-vitiating.[8]

This characterization does not enable us in general to distinguish between what we would take to be a true counterfactual in constrast to a false one such as the one containing 'h is made of blue cheese' as antecedent.[9]

Now the problem of determining the criteria for discrimination between true counterfactuals and false ones is not "a fussy little grammatical exercise."[10] expecially in the case we are considering. A major difficulty concerning the predicate 'self-vitiatable' lies in the fact that it appears at present in no recognized and well-confirmed body of theory. The suggestion, that a true counterfactual is one which rests upon

---

[8] It should be remarked that although our pre-systematic characterization of the phenomena which is described as self-vitiating involved the use of a counterfactual, our definition of 'self-vitiating' used only indicative assertions.

[9] As Goodman has shown. See Fact, Fiction and Forecast, Chapter 1.

[10] Goodman, op. cit., p. 13.

a well-confirmed law-like connection between its antecedent and consequent, will here not help us in the least. For the fact is that there is available no set of law statements at present from which, in regard to any particular h, we can deduce the conclusion 'h is self-vitiating'.

These assertions, about the state of our current empirical knowledge, are made by way of showing why it is not possible, at present at least, to clearly specify the criteria for application of the term 'self-vitiatable'. Rather than defining 'self-vitiatable' we must content ourselves with showing what steps must be taken in order to arrive at the kind of definition we are seeking. Fortunately, we do not need to wait, as Goodman has shown,[11] for a resolution of the general problems of disposition terms before undertaking the task of defining specific disposition predicates. The task before us is illuminated by Goodman's example of the disposition term 'flexible'. He writes:

> In dealing with a particular disposition, say flexibility, we may start with such predicates as 'bends' and '(is) under suitable pressure'. If both apply at one time, then the predicate 'bends under suitable pressure' applies; while if 'under suitable pressure' applies when 'bends' does not, then the predicate 'fails to bend under suitable pressure' applies.... We may also hereafter abbreviate 'bends under suitable pressure' as 'flexes' and 'fails to bend under suitable pressure' as 'fails to flex'.
>
> Now 'flexes' and 'fails to flex' are mutually exclusive, and together they exhaust the realm of

---

[11] Op. cit., pp. 44ff.

> things that are under suitable pressure; but neither applies to anything outside that realm. Thus from the fact that 'flexes' does not apply to a thing, we cannot in general infer that 'fails to flex' does apply. Within the realm of things under suitable pressure, however, the two predicates not only effect a dichotomy but coincide exactly with 'flexible' and 'inflexible'. What the dispositional predicates do is, so to speak, to project this dichotomy to a wider or even to the universal class of things; and a predicate like 'flexible' may thus be regarded as an expansion or projection of a predicate like 'flexes'. ...
>
> We can define 'flexible' if we find an auxiliary manifest predicate that is suitably related to 'flexes' through "causal" principles or laws. The problem of dispositions is to define the nature of the connection involved here: the problem of characterizing a relation such that if the initial manifest predicate '**Q**' stands in this relation to another manifest predicate '**A**' then '**A**' may be equated with the dispositional counterpart -- '**Q**-able' or '**Q**d' -- of the predicate '**Q**'.[12]

Goodman's treatment of 'flexes' offers us a clear method of proceeding. The two predicates with which we are concerned are 'uttered under suitable circumstances' and '(is) self-vitiating'. If both of these apply to a given utterance, then the predicate 'is self-vitiating when uttered under suitable circumstances' also applies. And in a case in which 'uttered under suitable circumstances' applies and 'is self-vitiating' does not apply, then likewise 'is not self-vitiating under suitable circumstances' applies.

The problem of defining 'self-vitiatable' is the problem of finding an auxiliary manifest predicate that

---

[12] Ibid., pp. 48-49.

is suitably related to 'is self-vitiating when uttered under suitable circumstances' through causal laws. One such law might take the following form:

$$(h)(Ah \cdot Bh \cdot Dh :\supset \cdot Ch \supset Vh)$$

where '**A**', '**B**', and '**D**' stand for certain specific characteristics of h and '**C**' stands for the predicate 'is uttered under suitable circumstances'. Provided this statement is unviolated, supported, and unexhausted, we have the means for defining the disposition predicate 'self-vitiatable'.[13] The definition would read as follows:

$$V^*h =_{df}. Ah \cdot Bh \cdot Dh$$

where '**V***' stands for 'is self-vitiatable'. Now provided that '**A**', '**B**', and '**D**' are themselves manifest predicates (or can be defined in terms of manifest predicates according to the method suggested by Goodman) we are then able to define 'self-vitiatable' without employing a dispositional and without reference to any such troublesome words as 'possible'or 'potentially'.

---

[13]The confirmation of such a statement raises other problems in confirmation theory; e.g. what constitutes violation and what constitutes support of a particular generalization? Although these problems would need to be considered in the development of any theory of self-vitiation they are not problems peculiar to our subject matter. We discussed certain other problems in confirmation theory in Chapters III and IV; (see also Goodman, Ibid., Passim.).

We have also eliminated the need for a counter-factual characterization of the phenomena of self-vitiatability. And the two predicates 'self-vitiating' and 'self-vitiatable' enable us to say what is cognitively significant about the phenomena in question; that is to say, we can find confirming evidence for the assertion '$\underline{h}$ is self-vitiatable' by determining whether or not the assertion 'Ah . Bh . Dh' is true.

We have remarked that the predicates apply only to actual utterances of $\underline{H}$ and not to possible ones. But there are surely occasions on which we would wish to speak of possible members of $\underline{H}$ having the characteristic of self-vitiatability. Suppose, for example, that $\underline{H}_1$ has only one member, $\underline{h}_1$, but this member does not have characteristics $\underline{A}$, $\underline{B}$, and $\underline{D}$. Now we may feel that we are justified in an assertion such as the following:

> If there were a member of $\underline{H}_1$, $\underline{h}_2$, such that $\underline{h}_2$ had characteristics $\underline{A}$, $\underline{B}$, $\underline{C}$, and $\underline{D}$, then $\underline{h}_2$ would be self-vitiating.

The justification for such an assertion would be its being supported by the law: (h)(Ah . Bh . Dh : ⊃ . Ch ⊃ Vh). But it may be noticed, the above counter-factual speaks of only possible members of $\underline{H}_1$ and not actual members of $\underline{H}_1$. And hence our definitions of 'self-vitiating' and 'self-vitiatable' would appear to be inadequate; for they do not allow us to speak of possible members of $\underline{H}$ but only actual members.

While the actual solution to this particular problem would undoubtedly require vast expenditures of time and energy in empirical research, the outline of the solution is again relatively clear from the foregoing analysis. The difference lies in the nature of the disposition in question. This can be made clear by rephrasing the counterfactual:

> There is a person p such that p has the disposition to utter self-vitiating replicas of h_1(sp).

Here the disposition term in question is not a predicate of possible, non-actual utterances but rather a predicate of actual persons. What we are suggesting is that the same method of analysis which has been used for the disposition terms 'flexible' and 'self-vitiatable' would be adequate for an analysis of the term 'has the disposition to utter self-vitiating replicas of h_1' and also the more general term 'has the disposition to utter self-vitiating hypotheses'. What would be required is the specification of certain characteristics of persons (or perhaps machines, computers, etc.) which, when present under suitable circumstances, lead by law to their uttering in those circumstances, self-vitiating hypotheses. And this empirical task would presumably be one for the psychologist or sociologist.

An extension of this method could take care of cases such as that in which we wish to speak of possible

persons uttering self-vitiating hypotheses. This case could be dealt with by specifying the conditions under which actual persons would have procreated in such a fashion as to produce such a possible person. Needless to say such an extension of method would be considerably difficult and complicated. But our point is simply that the method offers no insuperable theoretical obstacles to saying all that we wish to say using only manifest predicates which have no hypostatized entities in their extensions.

Up to this point we have been dealing with two classes of utterances: those which actually are self-vitiating and those which although not self-vitiating are self-vitiatable. Now there are a number of utterances which seem to "hover" between being self-vitiating and being self-vitiatable though not self-vitiating. An example may make this clear. Suppose that a scientist makes a fore-diction during 1962 concerning an event, e, which is supposed to occur during the year 2000 A.D. The fore-diction is not published at the time that it is made but is discovered in the effects of the scientist when he dies in the year 1990. The fore-diction is then published in 1990. Another scientist observes that the publication of these replicas of the sentential parts of the original fore-diction has the characteristics A, B, and D. The question then is: "Shall the initial utterance of the fore-diction be called 'self-vitiating' or

'self-vitiatable'?" The answer to this question will depend on whether or not the predicate '_C_' is applicable to the utterance. But here we note that a certain ambiguity occurs in our usage of the term 'utterance'. We have not clearly specified whether the sign-event which we have called 'an utterance' is to be construed as simply the _production_ of tokens or whether it is to be construed as the production and reception of tokens.

Let us, in the former case, speak of the utterance as being incomplete and, in the latter case, of a complete utterance. In an incomplete utterance the tokens have not been, so to speak, received. If the tokens have been received (either read, heard, felt, etc.) then the utterance will be called complete.[14] We suggest that the predicate 'self-vitiating' be applied only to complete utterances. We are assuming that prior to 1990 the fore-diction mentioned above was part of an incomplete utterance. But the question of its _eventual_ completeness is what is

---

[14]Leonard draws the distinction between undirected and directed sign-events on the basis of there being an intended second party or being no such intended second party. The distinction is inadequate for our purposes. Some sign-events, although having an intended second party will be incomplete in our sense (e.g. letters addressed incorrectly and eventually winding up in the dead letter office of the U.S. Post Office.). The question of eventual completeness is what is important for our analysis and not the quite different question as to whether the productive sign-event has an intended second party. Similar remarks obtain in regard to receptive sign-events. By a complete utterance we mean one in which some second party actually receives the tokens whether this second party is the producer or not. The only events then which are to be classed as incomplete are those in which the tokens are not received. See Leonard, _Principles of Right Reason_, pp. 89-107.

at issue here. And this issue presumably can, theoretically at least, be resolved by law. In fact the only time in which incomplete utterances are of any interest to us would be in those cases such as the fore-diction mentioned. For once the tokens are read or heard or received in some other manner, the utterance is, by definition, complete. And once it has been shown to be complete, whether or not it is self-vitiating will depend upon the causal consequences of its having been completely uttered, that is, produced and received.

Thus to call such an utterance 'self-vitiating' is to say that the utterance either did have, does now have, or will have a self-vitiating consequence. In other words, 'self-vitiating' is a predicate which is as tenseless in its application as 'is true'.

Interestingly enough, the predicates 'self-vitiating' and 'self-vitiatable' differ in some important respects from predicates such as 'is magnetic'. For example, of a particular iron bar we may say that it is magnetic or that it attracts iron filings at time t. We may also say that while it was magnetic it is magnetic no longer.[15] But it will not make sense to say that

---

[15] One of the principle reasons for the difference between 'is magnetic' and 'is self-vitiatable' is that in regard to the former we have no special term which is its manifest correlate or counterpart. So we use the same term for both the manifest property and the disposition. Our point, however, has to do only with the dispositional use of the term 'is magnetic'.

while h was self-vitiatable, it is no longer self-vitiatable, and especially so when h is a generalization. One of the characteristics of the sentential parts of an utterance of a generalization is that replicas of the generalization may be freely reproduced. Hence the initial utterance will enter into a causal sequence such that many ensuing utterances may become instantiations of some 'x_i' in our definition. And each replica, provided it has the characteristics A, B, and D, will also be construed as self-vitiatable.

In the case of singular statements certain of the replicas will presumably not be construed as self-vitiatable; for example, those utterances produced after ~e has occurred. But law-like generalizations will not be so restricted in regards to their replicas.[16]

In regard to self-vitiating hypotheses we may cogently apply the predicate to those which have proved to be self-vitiating or to those which according to the laws exemplified in the expressions 'Cx_ix_j', we may predict will cause ~e. This is not to suggest that our predictions may not sometimes turn out unsuccessful.

---

[16] Non-law-like generalizations may also be so restricted both in regard to 'self-vitiatable' and in regard to 'self-fulfillable'. Prima facie, the restriction seems to be dependent upon the relationship of h to e. Where the occurrence of e may be said to "completely" confirm h then if h is a generalization it is non-law-like. After e has occurred, any future utterance of H will presumably not be causally efficacious in bringing about e. See Chapter IV.

Any initial theory of self-vitiation, we may expect, will need to be continually revised prior to its being able to give us a preponderance of successful predictions. The point is that we will be justified in applying the predicate 'self-vitiating' to utterances both which have been "proven" to bring about ~e in the sense explicated, or which according to well-confirmed laws may be predicted to bring about ~e. The point here is one which is directly analogous to the Hempel-Oppenheim paradigm of scientific explanation. We do not distinguish logically between explanation and prediction. In like manner, when we say "h is self-vitiating" we are not saying whether the event ~e is past, present, or future. The sense of our claim that h is self-vitiating is that its occurrence leads by law to ~e.

We need also to make clear some of the relationships which obtain between the predicates 'self-vitiating' and 'self-vitiatable'. If h is self-vitiating, then it will also be self-vitiatable. And should a particular h prove to be self-vitiating without at the same time being self-vitiatable, then presumably something is wrong with the choice of the set of characteristics, A, B, and D. However, what would presumably have happened in such a case is that h failed to exhibit one or more of these characteristics although it actually is self-vitiating. What may then be needed is a revision of the set of characteristics which have been used as the defining set for the term 'self-vitiatable'.

It is also clear that not every utterance which is self-vitiatable will be self-vitiating. The difference lies in the fact that while a particular utterance h may have characteristics A, B, and D it will fail to exhibit characteristic C (i.e. being uttered under suitable circumstances) and hence will not, according to well-confirmed law-statements, be self-vitiating.

Pervasiveness of Self-vitiation and Self-vitiatability

The claim that the phenomena of self-vitiation is one which enables us to distinguish between the social sciences and the non-social or physical sciences is apparently untenable. What we have tried to show is that the phenomena in question are pervasive of all the sciences and are not limited to any one set of sciences. There are important distinctions yet to be made; but the distinction here in question is not one of them.

The phenomena of self-vitiation may be treated as a sub-class of two other classes of phenomena. First, it may be treated as a sub-class of what are sometimes called "perturbance effects". Second, it may be treated as a sub-class of phenomena such that predictions of their occurrence sometimes fail due to the fact that certain auxiliary assumptions regarding boundary conditions have not been made explicit. Certain phenomena in the physical sciences exhibit effects similar to those of the phenomena of self-vitiation. It is a well-known

fact that the measurement of the position and velocity of atomic particles creates a state of affairs such that the simultaneous measurement of both position and velocity does not seem to be possible according to present measuring techniques and theoretical physics. The actual observations themselves cause the particles to react in ways in which they would not react if they were not being observed.

A similar phenomenon occurs in social science. A sociologist may, in observing the behavior of a small group, G, unwittingly introduce a disturbing factor into the group, and the members of G may then act in a manner different from that in which they would act if they were not being observed.

These two examples tend to show that the kinds of disturbance one encounters in the physical sciences and in the social sciences in measuring certain features of the world do not exhibit differences which prima facie enable us to make relevant distinctions between the two kinds of sciences. We are contending that self-vitiation can be treated as a specific kind of problem whose resolution may in some respects be hastened by the development of a general theory (in the sense of a systematic philosophical analysis) of perturbance. Whether or not the contention is justified is a question whose answer will need to wait upon the development of such a general theory. But there do not appear to be reasons why such would not be the case.

In regard to the second point, self-vitiation may also be treated as a sub-class of those phenomena such that the predictions of their occurrence fail to be successful due to a failure of the prediction to include certain auxiliary assumptions regarding boundary conditions. Self-vitiating hypotheses also exhibit this particular characteristic. But whereas in the case of physics we are frequently able to specify the unaccounted-for boundary conditions after the failure of the prediction, in the case of self-vitiation in the social sciences we are by no means so fortunate. Again, what is required is a scientific theory of self-vitiation which will enable us in each case to specify the value of each '$\underline{x}_i$' in regard to a particular utterance. But this distinction is irrelevant to our thesis. It amounts simply to saying that in the one case (that of physics) we have a more adequate theory than we do in the other case (that of psychology and sociology). It is the same as saying that the task of constructing a theory of self-vitiation remains. And this is precisely what we have contended throughout the course of this essay.

On the other hand, we need remark that the extension of the term 'self-vitiatable' is not so broad as to be otiose. At least two major points may be noted. First there are a large number of fore-dictions in celestial mechanics which would not be construed as self-vitiatable due to laws of relativity physics as presently understood. A scientist $\underline{S}$ may utter a fore-diction $\underline{h}$

at time t to the effect that an event e will occur in the vicinity of the constellation Andromeda at time t_n. If we assume that the time required for light to travel from S to Andromeda is greater than n, then h will not, according to any adequate theory of self-vitiation, be construed as self-vitiatable. This is due to the fact that presumably the speed of light is a limiting velocity in the universe; and no matter how the tokens of h were transmitted they could not arrive in time to have a causal effect on e.

Second, it seems quite likely that an adequate theory of self-vitiation will include as defining characteristics of 'self-vitiatable utterance' a set of characteristics such that they apply to only a small percentage of actual utterances. We may note presystematically that, in regard to our example of disrupting the orbit of Mars, no actual utterances concerning Mars' behavior would be construed as self-vitiatable. In the social sciences presumably a much larger class of utterances would be self-vitiatable. But even in the social sciences we may expect the predicate to apply to only a small percentage of actual utterances. The actual extension of the term, however, will in part depend on the conventions employed for defining the term 'self-vitiatable'. On page 137 we suggested that a law of self-vitiation might take the form:

$$(h)(Ah \,.\, Bh \,.\, Dh :\supset .\, Ch \supset Vh)$$

And from this analysis we suggested that 'self-vitiatable' could be defined in terms of the predicates 'A', 'B' and 'D'. Now we may easily see that the above statement is logically equivalent to

(h)(Ah . Bh : ⊃ : Dh . ⊃ . Ch ⊃ Vh)

And if the law were formulated in this latter form, 'self-vitiatable' could be defined in terms of the predicates 'A' and 'B'. And since the inverse relation between extension and intension would presumably hold in such a case, the extension of the alternative definition of 'self-vitiatable' would be greater than would the extension of the former. We are not claiming that such would necessarily be the case. We only use this example to show that the choice of predicates for the definiens of 'self-vitiatable' could vary according to the purposes of the definer. And as the purposes varied, so also could very the extension of the term 'self-vitiatable'. At any rate it seems clear that the extension of the term would not necessarily be so broad as to apply to every utterance or even to a major proportion of all utterances.

Goal Directed Phenomena and Self-vitiation

We have now reached the point in our analysis where we can make distinctions which may prove useful in the development of an adequate theory of self-vitiation. The first of these distinctions has to do with goal-directed phenomena.[17] Certain physical or biological

---

[17]The treatment of goal-directed phenomena is dependent upon Nagel's analysis in The Structure of Science, pp. 401-428.

systems "exhibit in varying degrees adaptive and regulative structures and activities" which enable them to preserve themselves in some relatively steady state or process of development in the face of environmental changes or changes in the behavior of one or more of the parts of the system. Thus, for example, the human body tends to preserve an internal temperature which ranges between 97° F. and 99° F. We may call such states "goal-states" or 'G-states' or 'property G'.

> Let S be some system, E its external environment, and G some state, property, or mode of behavior that S possesses or is capable of possessing under suitable conditions. Assume for the moment (this assumption will eventually be relaxed) that E remains constant in all relevant respects, so that its influence upon the occurrence of G in S may be ignored. Suppose also that S is analyzable into a structure of parts or processes, such that the activities of a certain number (possibly all) of them are causally relevant for the occurrence of G. For the sake of simplicity, assume that there are just three parts, each capable of being in one of several distinct conditions or states. The state of each part at any given time will be represented by the predicates 'Ax', 'By', and 'Cz'. respectively, with numerical values of the subscripts to indicate the different particular states of the corresponding parts. Accordingly, 'Ax', 'By', and 'Cz' are state variables, though they are not necessarily numerical variables since numerical measures may not be available for representing the states of the parts; and the state of S that is causally relevant to G at any given time will thus be expressed by a specialization of the matrix '(AxByCz)'. ...
>
> One further important general assumption must also be made explicit. Each of the state variables can be assigned any particular "value" to characterize a state, provided the value is compatible

> with the known character of the part of S whose state the variable represents.[18]

We assume that S is a deterministic system; and also that the value of each state variable at any given instant is independent of the value of each of the other state variables at that same instant. Only certain values of the state variables will presumably be causally effective in bringing about a G-state in S. The parts of S, however, are so related that a change in one of the state variables causes a change in the other state variables. Thus if a change in S occurs such that the value of Ax deviates to the point that S is no longer in a G-state, S is so organized that the values of By and Cz change to the point that S returns at a later moment to a G-state.

> The prima facie distinctive character of so-called "goal-directed" or teleological systems is thus formulated by the stated conditions for a directively organized system... What may be called the "degree of directive organization" of a system, or perhaps the "degree of persistence" of some trait of a system, can also be made explicit in terms of the above analysis.[19]

By relaxing the restrictions on E and allowing it to vary we simply add additional variables to the

---

[18] Nagel, Op. Cit., pp. 411-412.

[19] Ibid., p. 417.

analysis and now consider a new system S' which is composed of S and E.

Our summary of Nagel's analysis of goal-directed systems outlined above will enable us to mark a distinction between two classes of utterances: those involving goal-directed receivers and those not involving goal-directed receivers. We believe that it is the former class which is of particular interest to us rather than the latter. The latter class includes such "far-fetched" examples as those we recorded earlier regarding icebergs, ice-cubes, and exploding cobalt-bombs. (Any uneasiness we may have felt in regard to including these examples in the extension of the term 'self-vitiating' may be accounted for by noting that we had not yet distinguished between those utterances involving goal-directed receivers and those which do not.) Henceforth in our discussion we shall limit ourselves to a consideration of only those utterances involving goal-directed receivers.[20]

Using the symbols above we may now attempt to analyze in some detail the process of self-vitiation. Let u be an utterer of an hypothesis h, r be the receiver of h, and S be the system composed of u, r, h(sp) and any other relevant factors in the environment of u, r, and h(sp). We assume that r is a goal-directed system

---

[20]Hence we do not feel obliged to coin a neologism to distinguish the two classes here considered.

and that h(sp) constitutes a fore-diction that e. Further, we assume that e constitutes a G-state of r. Now in the event that h should turn out to be self-vitiating, our problem is that of explaining this fact. An explanation would consist in showing how u in uttering h(sp) causes certain changes in one or more of the state variables of r. The result of these changes in the values of the state variables is that ~e occurs. In other words r is taken out of a G-state.

Now the analysis becomes more complicated when we note that systems such as r (e.g. if r were a human being) may be considered as exhibiting "self-regulatory behavior with respect to several G's at the same time, alternative (and even incompatible) G's at different times, a set of G's constituting a hierarchy on the basis of some postulated scale of 'relative importance' or more generally a set of G's whose membership changes with time and circumstance".[21] Now the difficult cases for a theory of self-vitiation would be those in which two incompatible G's appear to be involved at the same time. Suppose that in the above example h(sp) constitutes a fore-diction that r will be in some G-state at time t. Let us call this G-state 'G_1'. Let us suppose also that ~e is a necessary condition for the occurrence of another G-state in r at t. Let us call this G-state 'G_2'. Now it is

---

[21] Nagel, op. cit., p. 416.

obvious that it is not possible for r to be in both G_1 and G_2 at t. Hence reference to the original h which predicts that r will be in G_1 must be inadequate as an explanation or prediction of the phenomenon in question.

We must now take note of a fundamental difficulty with our characterization of what shall constitute a self-vitiating hypothesis. The difficulty involves requiring that h must be deducible according to the requirements outlined in Chapter I, i.e. from a set of *true* statements. What the foregoing shows is that this requirement must be relaxed to read:

> h must be deducible from a set of statements which we would be warranted in believing to be true prior to the determination of h's being self-vitiating.

As we pointed out earlier,[22] unless we make this modification the extension of the term 'self-vitiating hypothesis' will be null.

We can now see that the theory on which the deduction of the claim that r will exhibit G_1 at t is faulty and that it is the self-vitiating character of h which exhibits the faultiness of the claim. In Chapter VI we shall examine some possible ways of deciding between two such incompatible predictions (e.g. between h and 'Vh').

---

[22] See footnote, page 64.

At this point we wish merely to clarify some of the difficulty involved in the event that a system r may exhibit incompatible G-states.

Nagel suggests a solution. We need to develop in the case of r and all like systems a list of "G's constituting a hierarchy on the basis of some postulated scale of 'relative importance'." In the case of our example the failure of h may be traced to the fact that some such scale of importance regarding the relation of G_1 to G_2 in r was not considered by u in the deduction of h. It may be expected that the more complex the system involved and the greater number of G-states to be considered the greater would be the difficulty in specifying the "relative importance" of the various G's. Again, we are not minimizing the magnitude of this empirical task but simply attempting to show that there are not theoretical obstacles in the way of developing and adequate theory of self-vitiation.

Sign-behavior and Self-vitiation

The above analysis effectively discriminates between the kinds of self-vitiation involved in the Mars' example and those involving goal-directed behavior. There is, however, another equally important distinction to be drawn. This distinction involves those goal-directed systems which exhibit what we may call "sign-behavior" and those which do not. It is perfectly possible that a self-vitiating utterance may involve a

receiver such that the receiver may be understood as a goal-directed system and yet not be one which exhibits sign-behavior.[23] The interesting cases involve not only utterance reception (as thus far construed) but sign-interpretation. It is this latter feature which creates the most puzzling (from a presystematic standpoint) and cognitively important instances of self-vitiation. In this section we shall not undertake to give anything like an adequate account of the phenomena of sign-interpretation but shall limit our discussion to what we consider to be the points most relevant to our problem.

To begin we may return to our previous characterization of what constitutes a replica. We introduced the notion of 'replica' in Chapter III as follows: "x is a replica of y if and only if x and y share certain specifiable physical characteristics." This characterization enabled us to explicate the concept of a self-vitiating utterance without having to take account of all the problems at once. Now we must note that this particular characterization is inadequate to our task. This inadequacy is due to the fact that two spoken utterances may indeed contain sentential parts which we would wish to include in a single h or H from a presystematic standpoint. Yet the physical characteristics

---

[23]See our example on p. 108, Chapter IV.

of these sentential parts might, from the standpoint of phonemic analysis, be quite unlike. For example, John F. Kennedy will, in his ideolect, use a quite different set of phonemes for the production of the word 'idea' than will Dwight D. Eisenhower. And the set of phonemes thus produced by Kennedy will probably share, in Eisenhower's ideolect,[24] more physical characteristics with 'I, Dear' than with 'idea'. This example serves to show that, while the notion of replicahood as so far construed is adequate perhaps for the written language, it is quite inadequate for the spoken language. And presumably we shall wish to include spoken as well as written utterances in a specific h. Due to the fact that great varieties of phonemic differences occur in different dialects and different ideolects, some means must be available for determining when a particular verbal utterance shall be included in a particular h.

Two solutions to this problem are immediately apparent for our purposes. The first is to specify all the relevant semantical and syntactical rules of the language. These would enable us to determine not only when a pair of utterances h_1 and h_2 are describing the

---

[24]The term 'ideolect' is taken from Paul Ziff, Semantic Analysis, (Ithaca, New York: Cornell University Press, c.d. 1960) who in turn has adopted the usage of U. Weinreich, "Is a Structural Dialectology Possible?". Linguistics Today, p. 269. It refers to the characteristic features of a language as that language is spoken by one particular person.

same event e, but also when they are or are not well-formed expressions. The determination of which utterances shall be members of a specific H on the basis of semantical and syntactical rules would undoubtedly prove to be an exceedingly complicated and, in most cases, difficult task. A second solution may prove to be more easily manageable. This solution is the one we indicated in Chapter III. It involves the introduction of the notion of an ideal observer. An ideal observer would presumably (for our purposes) have the following (among other) characteristics: a) be able to read and write the language L; b) have a "trained ear" so as to be able to "understand" the various ideolects and dialects of L; and c) be absolutely fastidious in spelling and punctuation. We can now deal with the problem indicated above on the basis of our second solution: if h_1(sp) is transcibed by an ideal observer O as h_3(sp) and h_2(sp) is translated, by the same (or possibly different) observer, as h_4(sp), and h_3(sp) and h_4(sp) are replicas of each other; then h_1(sp) is a replica of h_2(sp) and conversely.

Our test does not eliminate the need for semantical and syntactical rules in determining whether a pair of utterances are members of some particular h. Our semantical rules will still be required in determining whether or not a specific h describes e; and syntactical rules will be required in determining whether or not a specific h is well-formed. But our test does eliminate the need

for these rules in determining whether or not verbal utterances are to be construed as replicas of one another.

It may be objected that our test is highly inadequate inasmuch as what will usually be involved is not a comparison of $\underline{h}_1$ and $\underline{h}_2$ (when these are spoken utterances) but a comparison of what is remembered as $\underline{h}_1$ and $\underline{h}_2$ or else a comparison of a new set of utterances $\underline{h}_5$ and $\underline{h}_6$ which are presumed to be replicas of $\underline{h}_1$ and $\underline{h}_2$ respectively. We seldom have ideal observers recording all those utterances or even a small percentage of those utterances which we would wish to consider as instances of self-vitiation. In most instances we shall be dependent upon someone's memory as to what was said and what was meant in regard to a specific utterance.

We may readily admit that in regard to spoken utterances this factor presents a formidable problem. But it is one which would attend any investigation of self-vitiation whether one used semantical and syntactical rules for determining the question of replicahood in any specific case or whether one used the test which we have proposed. Our test does not circumvent this particular problem. But it does provide an effective means for

circumventing other problems in linguistic analysis and hence simplifies the task at hand.[25]

The preceding considerations concerning phonesis point up a difficulty considered earlier in this essay. In regard to utterances which are vague, ambiguous, or mal-formed we decided to treat these as values of some 'x_i' in our definition rather than as members of h. We must now consider ways of dealing with those utterances such that the receivers do not understand or else misunderstand what is signified by the sentential parts of the utterances. These utterances are complete, but due to ideolectical difference between producer and receiver (or other factors) the receiver does not understand or misunderstands what is indicated by the utterance. A number of kinds of circumstances could bring about such a state of affairs. An utterer u could stutter, lisp, or perhaps have suffered a tracheotomy; or he could speak the language in a dialect which differed so greatly from that of the receiver r that r would be unable to determine what is indicated by the set of phonemes employed by u. Yet an ideal observer O presumably would be able to determine that which is indicated by u's tokens. On the

---

[25]The effectiveness of the test of course will depend on the linguistic ability of the observer. We have not done away with the need for the rules in question; we have simply allowed ourselves to trust the judgment of a presumed authority.

other hand such utterances may turn out to be self-vitiating, even though the receiver of the tokens did not understand what was indicated by the tokens.

The way in which we may deal with these utterances (which to distinguish them from h* utterances we shall call 'h** utterances') will depend on certain other circumstances. Three cases may be distinguished: 1) h**(sp) is a replica of h(sp) and h is uttered prior to h**. Further, h is a cause of h** as well as ~e. Provided that h** enters the causal sequence eventuating in ~e, it may be treated as a value of an 'x_i'. In this case the predicate 'self-vitiating' would apply to h but not to h**. In this case we assume that the important difference between h and h** is that the former is understood by the receiver of h whereas the latter is not understood by the receiver of h**. 2) In this case the occurrence of h** precedes that of h, and is a cause of h. Provided that h is a cause of ~e, h** may be ignored and the predicate 'self-vitiating' applied to h. 3) In this case there is no h which is a replica of h** and which leads by law to ~e. Yet h** does lead by law to ~e. In this case the predicate 'self-vitiating' would be applied to h**.

Now the third case, even though it involves goal-directed receivers will be treated as a somewhat uninteresting case of self-vitiation. Inasmuch as the tokens of h** are not understood by the receivers of these

utterances (i.e. the h** type as in 3), this class may be dealt with in a manner similar to that of the raven and Mars examples.

The cases such as those in 3) do, however, point up the characteristic feature of the kind of self-vitiating hypothesis which is cognitively important inasmuch as this feature is lacking in utterances such as those considered in 3). This feature may be described as follows: the receiver r of tokens h(sp) produced by u understands the semantic relations which obtain between h(sp) and the state of affairs which would be the case if h were true. The term 'understand' used in the above description may prove to be subject to the kind of ambiguity discussed in relation to the term 'know' discussed in Chapter II. But this problem need not detain us. We may use the term 'understand' in the following sense: "r will be said to understand the semantic relations which hold between h and that which would be the case if h were true if and only if r believes that the statement 'Fh ≡ ~e' is true. And further, the assertion '"Fh ≡ ~e" is true' is true. Less formally, it is not only the case that r believes that certain semantic relations hold between h and that which would be the case if h were true, but that such semantic relations *do in fact* hold.[26]

---

[26]We have once again been forced to resort to the counterfactual form in order to speak of the semantic relations concerning h. We are assuming, however, that the same mode of analysis used above would be adequate for

Now it may appear that our classification of self-vitiating hypotheses will enable us to discriminate between the social and non-social (or physical) sciences due to the fact that we have introduced the notion of 'belief' as a characteristic of receivers of the class of utterances here considered. Believing we usually take to be a characteristic of sentient and rational creatures. But the term 'belief' need not be so narrowly construed. We may (paraphrasing Leonard) define '_x_ believes _y_' as '_x_ has a disposition to act as though _y_ were true'.[27] And Nagel's example quoted in Chapter II shows that certain mechanical devices may have such dispositions. Applying the term 'belief' to machines may seem to strain customary speech. But we are not primarily interested in defending "ordinary language" although such a defense may at times be warranted. And we may always avoid any seemingly unconventional "connotation" which the term 'belief' has in such contexts by employing the expression in the definiens of the definition of the term 'belief', namely, '_x_ has a disposition to act as though _y_ were true'.

This digression should not obscure the fact that the belief in question is not belief in the truth of _h_ (although in some instances this fact also may be

---

disposing of the problems encountered in this regard. This would obviate any need for using an expression '_e_' when '_e_' does not denote.

[27]See Principles of Right Reason, p. 46.

important) but in the truth of the assertion that certain semantic relations hold between h and what would be the case if h were true. In a number of cases the receivers may actually attempt to act in such a way that h turns out to be false.[28]

The simplest instances of the class of self-vitiating hypotheses here being considered are those in which what h purports to describe is such that a receiver r would be taken out of a G-state if what h purports to describe were in fact the case. If r is a goal-directed receiver then, provided other things are equal, we may presume that r will act in such a way as to attempt to bring about ~e. We shall, of course, need to stipulate that r's attempts will have some probability of success, even though this may in some cases be small. But in this case of self-vitiation, we are here considering that what occurs is a success by r in bringing about ~e.

---

[28] A comprehensive theory of self-vitiation would need to consider certain empirical questions such as the manner in which human beings learn to speak a particular language L, and how they learn to discriminate between various homonyms. Linguistic theory may be considered to have a relation to a theory of self-vitiation somewhat similar to that which obtains between a theory of physics and chemistry on the one hand and biology on the other. Though the biologist is dependent upon the findings of physics and chemistry, it is not required that he become a physicist or chemist in the pursuit of his special field. In like manner, one involved in the development of a theory of self-vitiation would not be required to solve all the problems in linguistic theory before undertaking his task, even though a solution to some of the problems in linguistic theory will aid in the solution to problems in a theory of self-vitiation.

More complicated situations may be dealt with taking into account certain considerations elaborated above. When more than one G-state is involved we may need a scale of "relative importance" before we can determine whether or not r will attempt (or to what extent r will attempt) to bring about ~e. If, for example the occurrence of that described by h would take r out of G_1, but efforts to bring about ~e would take r out of G_2, and G_2 (according to some scale of "relative importance",) is more important than G_1, then presumably r will not (provided r understands certain of the causal consequences of his action and provided he acts in a rational manner) attempt to prevent e from occurring.[29]

In sum, we are suggesting that the most important class of self-vitiating hypotheses are those which involve goal-directed receivers at least one of whose G-states would be displaced by another state (not a G-state) if h were true.

Where a given h involves a number of receivers it may be the case that a large number of G-states may be also involved. In some instances these may even be mutually incompatible inasmuch as it may be physically "impossible" for receiver r_i to remain in G_1 while receiver r_j remains in G_2.

---

[29]What is needed at this point is an empirical theory of value; i.e. a theory which would enable us to predict in any given instance how men actually choose when confronted with a "conflict of values". We have suggested elsewhere in this essay how normative questions of value, i.e. how men ought to choose, are also important

All of the foregoing remarks lead us to make two observations. First there are no theoretical obstacles to developing a theory of self-vitiation. On the other hand, the actual construction of such a theory will depend on the availability of a vast array of evidence and on the confirmation of many laws which are not yet available to us. There are no insuperable obstacles to the development of such a theory; and yet the obstacles which do remain are patently formidable.[30] This contention is underlined by Popper's arguments against historicism and historicists; even though we cannot agree with his arguments that knowledge of the future is logically impossible, we can certainly agree that at present we have very little in the way of acceptable theory which would justify any present claims that we can now predict, with any probability of success, how our fellows will behave in certain situations. The import of our remarks is to the effect that without an adequate empirical theory of self-vitiation the social scientist has little chance for anything but meager success in predicting what the morrow will bring. That the physical

---

in regard to the phenomena of self-vitiation. Normative questions, however, are not at issue here.

[30]We may secure some idea of the magnitude of the task if we attempt to discover all of the relevant G-states of human beings and then construct a scale of relative importance by which we could determine in a given set of circumstances how any individual receiver of tokens would act. It is belaboring the obvious to remark that a great deal of work yet needs to be done.

scientist is not plagued by this problem to as great an extent seems due to the fact that in most cases he is not required to predict the behavior of goal-directed receivers who can understand the fore-dictions which he utters. But as more and more use is made of such things as computers and electronic brains, we may expect that the physicist also will become increasingly concerned with this problem.

Summary

In this chapter we have attempted to explicate the sense of the term 'self-vitiatable' and then to show why it is not possible at present to clearly define this particular disposition predicate. We then went on to show how the cases of self-vitiation which are cognitively significant are those involving goal-directed receivers and especially those receivers which have the disposition to interpret signs. Neither of the two distinctions we considered, however, enable us to discriminate methodologically between the social and physical sciences. And we hope that this particular argument may be safely laid to rest.

In the next chapter we shall re-consider some of the arguments considered earlier (in Chapter II); we shall attempt to show in a more precise manner than was formerly possible that these arguments lack the cogency usually attributed to them.

CHAPTER VI

RECAPITULATION

Introduction

In this chapter we wish to re-examine, in the light of our last results, some of the arguments in previous chapters (especially Chapter II), review the difficulties they involved and suggest some solutions to these difficulties. We shall turn our attention primarily to the arguments put forward by Gewirth and Popper.

Does Self-vitiation Make Knowledge of the Future Impossible?

We recall that Popper claims that successful prediction of the future is impossible due to the fact that our knowledge of the future is dependent on knowledge of future knowledge and that this is logically impossible. Now we wish to make his point even more forcefully before proceeding to show that there are ways of circumventing the difficulty it apparently presents. Let:

'$\underline{T}$'	stand for	$\underline{L}_1 \ldots \underline{L}_n$ conjoined to $\underline{C}_1 \ldots \underline{C}_n$ (where any $\underline{L}_i$ is a general law-statement characterizing self-vitiating hypotheses and any $\underline{C}_i$ a statement describing the relevant antecedent conditions concerning a particular hypothesis $\underline{h}$.)

Then according to some theory of self-vitiation:

$$G_{(T,h)} \longrightarrow Vh$$

where '*G*' is a second degree predicate which holds between the conjunction of law-statements and antecedent-condition statements on the one hand and utterances on the other.

In other words, when the law-statements are taken in conjunction with certain statements describing antecedent conditions the conjunction of these statements logically implies the statement that *h* is self-vitiating. The fore-diction in question then is the statement '*Vh*'.

Now suppose that a social scientist *D* publishes a fore-diction regarding a particular utterance h_i to the effect that h_i is self-vitiating. Let us call this utterance of *D* at time *t* 'h_k'. This utterance h_k contains as a sentential part a replica of the tokens 'h_i is self-vitiating'.

Suppose further that *D* publishes his fore-diction h_k. We can significantly inquire whether his utterance 'h_i is self-vitiating' is *itself* self-vitiating. We may now note some of the advantages of taking utterances rather than tokens, or propositions, etc., as values of our variables. For the kind of paradoxes which are generated in such instances as the Russellian or the Grelling do not arise in regard to the predicate 'self-vitiating'. The values of the variables are not tokens but utterances which may be characterized by employing replicas of the tokens in question *but need not be*. We can always avoid any self-referential feature by giving

names to these utterances which in no way resemble the sentential parts of the utterances in question. Hence we need no recourse to a theory of types in order to deal adequately with the phenomena in question.

The question then as to whether or not h_k is self-vitiating is one which may be answered by an appeal to empirical evidence and to the laws of self-vitiation. If h_k is self-vitiating then h_k is false; but this then means that at least one of the L_i's or one of the C_i's is false. This is due to the following considerations: if Fh_k then it is false that Vh_i; and if 'Vh_i' is false, then '$G_{(T,h_i)}$' is false; but then at least one of the statements composing T must be false.

In such a case we would say that our theory of self-vitiation is inadequate. A more serious difficulty arises, however, in the following circumstances. Suppose that scientist E on the basis of the same theory of self-vitiation used by scientist D together with other antecedent condition statements predicts that the utterance of scientist D is self-vitiating. That is, he is asserting that Vh_k. Now the difficulty is as follows: insofar as two fore-dictions are related as described above, it will never be logically possible for both of these assertions to be true. For if 'Vh_i' is true then by definition of 'V' 'Vh_k' is false since the sentential part of h_k is a replica of 'Vh_i'. And if 'Vh_i' is true then it cannot be self-vitiating. If, on the other hand, 'Vh_k' is true then

it is false that $\underline{Vh_i}$. For if '$\underline{Vh_i}$' is self-vitiating, then it is false.

An example may make these formal points easier to keep in mind.

1. The National Safety Council publishes a fore-diction to the effect that four hundred Americans will die on Hd. Let us call the individual composed of these utterances '$\underline{h_o}$'.

2. Scientist D publishes a fore-diction on the basis of a theory of self-vitiation T that $\underline{h_o}$ is self-vitiating. Let us call the individual composed of certain of these utterances $\underline{h_m}$. (We may note that the sentential parts of these utterances will be of the form '$\underline{Vh_o}$'.)

3. Scientist E publishes a fore-diction on the basis of the same theory T that $\underline{h_m}$ is self-vitiating. Let us call the individual composed of these utterances '$\underline{h_n}$'. The sentential parts of these utterances will be symbolized '$\underline{Vh_m}$'.

Now since all the above fore-dictions have been deduced from well confirmed law-statements and highly evidenced antecedent condition statements we may inquire whether we have grounds for choice between them. We cannot hold all three of them to be true. While 1 and 3 can both be true, 2 and 3 cannot both be true nor can 1 and 2 both be true. Fortunately, there <u>are</u> grounds for choosing between them. One such set of grounds has to do with the nature of the auxiliary assumptions and the

boundary conditions discussed in earlier chapters. On this basis we would choose 2 in preference to 1 inasmuch as the explanans of 2 contains assertions about relevant boundary conditions not contained in 1. Likewise (and this is where the presumed difficulty lies) we would choose 3 in preference to 2 since 3 makes certain other boundary conditions explicit while 2 does not include statements concerning these boundary conditions although it does contain statements not included in 1.

It becomes immediately obvious that this "schema" enables us to produce a rather lengthy sequence. For example:

4. Scientist F predicts that Vh_n.

In this case we would choose 4 over 3 since it takes into account boundary conditions not described in 3, namely the results of the publications of tokens which are replicas of '$\underline{Vh_m}$'. What this boils down to is the question of which of the fore-dictions in question (i.e. 1, 2, 3, or 4) is better evidenced. And here important value considerations intrude themselves. We need to be able to specify when a particular fore-diction is better evidenced than another. And we would need this criterion before we could choose between 1, 2, 3, and 4. Also it seems intuitively clear that simply the addition of statements describing other boundary conditions is not sufficient grounds for choosing a higher number in the

sequence unless we are prepared to assert that each of the statements describing the antecedent conditions in n is as well-evidenced as each of the statements in n-1.

It may appear that the "schema" described above will always lead us into an infinite regress and hence any prediction of the future will never have any degree of certainty. Two points are at issue here. First it is not the case that such a schema leads to an infinite regress. Predictions and publications of fore-dictions are the kinds of things that are carried out in time. And prior to the occurrence of any predicted event e (or ~e), there will only be a finite amount of time available for the filling out of the schema. Hence the number of fore-dictions about fore-dictions ... about fore-dictions will always be finite.

We must not suppose, however, that the last fore-diction in the sequence will be supported by better and more conclusive evidence than the others. Being the last member in the sequence then is not a sufficient condition for our accepting it over any of the prior ones in the sequence.

Also, there might well be certain psychological limitations involved although their exact nature might be a difficult problem for a present-day psychologist to resolve. The limitation may be illustrated as follows:

Publication of 1 would presumably have some causal effect. Publication of 2 could have a much smaller or less important effect.[1] Publication of 3, since it might presumably be of interest only to a small coterie of social scientists would have little observable effect on the population as a whole in the relevant time span. If some such psychological limitation is indeed present then it would seem that the number of predictions needed in any actual sequence for a successful prediction would be small. A general theory of self-vitiation presumably would take cognizance of this fact and would be substantiated by an examination of instances of self-vitiation on a number of "levels".

The second point at issue concerns our degree of certainty or confidence that a theory of self-vitiation will give us nothing but successful predictions. If what is required of a theory before it is accepted is that it come attached with some kind of built-in guarantee that it will give us only successful predictions about the future, then obviously we will accept no theory whatsoever. To expect certainty is patently to misconstrue the nature of science. Moreover, no theory at present can take into consideration all the possible boundary conditions without considering, e.g., every possible

---

[1] In any of the pertinent senses of 'importance' explicated by Nagel, op. cit., pp. 582 ff.

configuration of atomic particles in the universe. Most of these configurations, however, are not relevant, or, at least, unlikely to be relevant or even to occur. What a theory, therefore, needs to account for are the relevant and probable boundary conditions. And a theory of self-vitiation will need to do only this and be well-confirmed to be acceptable. If there is any noticeable difference between, say, a theory of physics and a theory of psychology it would be that there are fewer relevant boundary conditions in physics (at least as presently understood) than in psychology which need to be taken into account in any actual prediction. But what constitutes a relevant boundary condition is, itself, not wholly clear. And until an adequate definition of 'relevance' is forthcoming, we cannot use relevant boundary conditions as a criterion for discriminating between the social and physical sciences.

At any rate it is clear that no theory offers us certainty. Whether or not we accept a theory of physics or a theory of self-vitiation will depend in large measure on how well-evidenced the theory is and not on whether it comes to us packaged with a label guaranteeing only successful predictions. In short, a theory of self-vitiation offers us only contingent knowledge of the world. To wish for more or to demand more of any theory seems to be a desire for that which no empirical theory

of self-vitiation can provide. It is something like requiring a round square and is based equally on a confusion about the object.

With the foregoing discussion as background we may now deal more fully with Popper's claim that knowledge of future knowledge is impossible. We have already seen that it is tautological to assert that future knowledge is not present knowledge and that insofar as Popper's thesis comes to this it is surely true -- but, as surely, quite trivially unportentious. However, interesting cases would be posed by those situations in which a fore-diction about future knowledge is actually self-vitiating.

Suppose scientist F predicts that in 2000 A.D. and not before the scientific community will possess a law K. Here 'K', we may assume, characterizes the law rather than merely names it or states it.[2]

Whether or not the fore-diction of scientist F would be self-vitiating could be determined by an adequate theory of self-vitiation. What would be involved here is the question as to whether the characterization of the

---

[2]These are the only interesting cases we need to consider. If 'K' names the law, then the only thing which the fore-diction is claiming is that there will be some law which will have that name. But it doesn't say what this law will be; i.e., it does not give the import of the law. If, on the other hand, the fore-diction states the law, then chances are that this case will fall under those of the tautological variety: i.e., present knowledge is not future knowledge and conversely. These comments are, of course, subject to the qualifications made in Chapter II.

law generated sufficient interest among scientists to motivate them to attempt to formulate and confirm the law characterized by 'K' in advance of 2000 A.D.

If this fore-diction should turn out to be self-vitiating, then a scientist G presumably could predict that the utterances would have such a consequence; moreover, by taking more of the relevant boundary conditionss into account G could more accurately predict when the scientific community would possess K (in some suitable sense of 'possess', e.g. have confirmed K to a high degree. See Chapter II, pp. 45-56).

Here we see that the same kind of sequence can obtain which we noted in regard to predictions of fatalities on Hol. The major difference between the two examples involves only the difference in the substance of what is predicted. In the latter case (i.e. the one concerning fatalities) the time isspecified and then the character of the event which is to take place at that time is predicted, whereas in the former case, the character of the event is specified and the time at which it will occur is predicted. In this instance, actually two predictions are being made simultaneously: 1) that the event will actually occur; and 2) that it will occur at such and such a time. But these differences should not obscure the fact that the kind of explanation or prediction which should prove adequate in the two cases will not significantly differ. In both cases what particular

prediction will constitute the final one in the sequence will be determined in part by the amount of time elapsing between the initial prediction and the event predicted. And whether we accept the final member in the sequence or some other as the most likely candidate for success will depend on how well-evidenced are the statements in the prediction and how many of the relevant boundary conditions are included in the explanans.

A Re-examination of the Arguments of Feuer and Gewirth

We recall that Gewirth and Feuer have made claims similar to those of Popper. Feuer claims that human creativity makes "for the unpredictable aspect of historical causation," and Gewirth has claimed that human decisions "may enter the antecedent or between the antecedent and consequent of laws". We have had occasion to point out the inadequacies of their analyses and the conclusions which they have drawn from them. But the kind of phenomena with which they are concerned is important for an adequate empirical theory of self-vitiation. We shall in this section make a case for their position as strong as possible, attempt to show precisely where the difficulty lies and then suggest some considerations which would be involved in a solution to these difficulties.

Suppose we now have a theory of self-vitiation. It is presumably adequate in the following senses: all of its statements are well supported, unexhausted (when

these are generalizations), and unviolated. Further, it has proved highly successful in predicting when any given utterance is self-vitiating. We assume also that the theory is able to account for the self-vitiating character of utterances on any of the "levels" of prediction discussed earlier.

Now let us suppose that armed with our theory we encounter a particularly obstreperous individual whom we shall name '$\underline{J}$'. $\underline{J}$ knows something about the logic of scientific explanation. On the other hand $\underline{J}$ intensely dislikes having us able to predict his behavior successfully. He feels that such prediction constitutes an unwarranted intrusion on his freedom. He therefore candidly tells us that if we predict that he will act in a manner described by an utterance $\underline{h}$, then he will act in a manner described by an utterance $\underline{\tilde{h}}$. And if we predict that he will act in a manner described by $\underline{\tilde{h}}$ then he will act in a manner described by $\underline{h}$.[3]

If our theory is adequate, it will have taken such individuals into account. What we presumably could deduce from such a theory is that under these circumstances

---

[3]Only certain values of $\underline{h}$ and $\underline{\tilde{h}}$ (here treated as variables) need to be considered. Utterances describing certain physical "impossibilities" are excluded from our consideration as uninteresting cases. For example if $\underline{h}$ contains as sentential part a sentence which is a replica of '$\underline{J}$ will jump over the moon at time $\underline{t}$ without the aid of mechanical devices' we shall not consider the pair of utterances $\underline{h}$ and $\underline{\tilde{h}}$ in this analysis. Our reason is simply that there does not appear to be any way in which such an utterance could be self-vitiating. What we wish to deal with are those cases where <u>both</u> alternatives lie within the realm of "possibility" for the actor $\underline{J}$.

that both $\underline{h}$ and $\underline{\tilde{h}}$ are self-vitiatable. Or we could deduce that the predictor $\underline{u}$ has a disposition to utter self-vitiating replicas of $\underline{h}(sp)$ or $\underline{\tilde{h}}(sp)$.

It should be noted that, although the sentential parts of $\underline{h}$ and $\underline{\tilde{h}}$ are of the form '$\underline{P}$' and '$\tilde{}\underline{P}$' respectively, the prediction in question is not a tautology (i.e. of the form '$P \vee \tilde{} P$') nor of a contradiction (i.e. of the form '$P \cdot \tilde{} P$'). This is so even though the sentential parts of $\underline{h}$ and $\underline{\tilde{h}}$ are logical contradictories. Our theory enables us to deduce:

$$V^{*}h \cdot V^{*}\underline{\tilde{h}}$$

And from this assertion it does not appear possible to deduce a contradiction. We recall that the definition of self-vitiatable will contain as definiens certain predicates (non-logical) which apply to sign-events. And there is nothing contradictory in asserting that both $\underline{h}$ and $\underline{\tilde{h}}$ have the characteristics $\underline{A}$, $\underline{B}$, and $\underline{D}$.

Our theory also, presumably, would enable us to deduce under these circumstances:

$$(Vh \vee V\underline{\tilde{h}}) \cdot \tilde{}(Vh \cdot V\underline{\tilde{h}})$$

That is to say, although both $\underline{H}$ and $\underline{\tilde{H}}$ contain a member which is self-vitiatable, only one of the pair contains a member which is self-vitiating. The first component of this conjunction is not a tautology. We recall that

the definiens of the definition of 'Vh' contains a number of empirical assertions. And it could well turn out that neither h nor $\tilde{h}$ is self-vitiating. Thus far our theory presents no logical difficulties. But the task of predicting J's behavior remains. If we take our task of predicting J's behavior as that of determining whether h or $\tilde{h}$ is true, then what our theory has been able to deduce up to this moment is inadequate for this task. What we wish to know is whether J will act in a manner described by h or in a manner described by $\tilde{h}$.

To make the requirements as stringent as possible we demand that whatever fore-diction is uttered by u must be understood (as explicated previously) by J. (It will turn out that this requirement constitutes the major difficulty in regard to successful predictions concerning this kind of self-vitiating utterances.)

Given the conditions specified above we might expect that our theory would enable us to deduce the following:

6.1 If u predicts h then J will act in a manner described by $\tilde{h}$.

6.2 If u predicts $\tilde{h}$ then J will act in a manner described by h.

This is to say that if u predicts h then h is self-vitiating and if u predicts $\tilde{h}$ then $\tilde{h}$ is self-vitiating.

At least two points need to be considered. First, we assume that u would not predict both h and $\tilde{h}$. What we require for a prediction is that the explanans contains statements which we are warranted in accepting as true.[4] And if u were able to deduce both h(sp) and $\tilde{h}$(sp) from a body of theory we assume that he would not be warranted in accepting all the statements in the explanans as true; and hence either one or the other or both of the utterances would not be considered a fore-diction. Second, the locution "u predicts h" is intended to mean that h is deducible according to the requirements for a fore-diction, and that h is a complete utterance in which J is the receiver. These requirements are already implicit in our use of the literal 'h'; therefore we could rephrase 6.1 and 6.2 as follows:

6.11 If u utters h then J will act in a manner described by $\tilde{h}$.

6.21 If u utters $\tilde{h}$ then J will act in a manner described by h.

Here we assume that while h and $\tilde{h}$ may be actually utterances their sentential parts will not both be deducible from the same body of theory.

So far our theory has told us nothing about J's behavior which he himself has not claimed to be the case.

---

[4]See p. 154.

What J has claimed is precisely what 6.11 and 6.21 state will follow as a consequence of certain utterances.

Suppose now that u has a body of psychological theory from which he can deduce h(sp). Let us call this theory 'T'. But he is also armed with a theory of self-vitiating T' from which he can deduce 6.11. He now knows that if he utters a replica of h(sp) (which he had deduced from his body of psychological theory) J will act in a manner described by $\tilde{h}$. On the basis of T' u concludes that T is inadequate. It is inadequate since presumably an utterance of a fore-diction deducible from T will be self-vitiating. Now u has made some strides in understanding, but he cannot yet predict J's behavior. All he knows is that if he utters h then $\tilde{h}$ is true. The same arguments apply *mutatis mutandis* to 6.21.

It may seem that we really do know how J will act in such circumstances and that our problem is a bogus one. After all, it may be argued, if I can deduce 6.11 from T' and u utters h, I can almost certainly successfully predict that J will act in a manner described by $\tilde{h}$. This fact should not be minimized. And while it does suggest that such a theory of self-vitiation would be extremely valuable, we have not yet come to grips with the central problem. The difficulty lies in what one takesto be the *fore-diction* which is used to predict J's behavior. In our example J is taking the fore-diction to consist of h since this is the utterance of which he

is the receiver. Our objector takes the fore-diction to consist of $\tilde{\underline{h}}$ since this is deducible from $\underline{T}'$ together with the assertion that $\underline{u}$ utters $\underline{h}$. But we cannot allow the solution to our problem to hinge upon a bald-faced ambiguity. Nor can we allow the solution (at least at this point) to depend upon deceit. If our fore-diction actually consists of $\tilde{\underline{h}}$ which we deduce from 6.21 plus the assertion '$\underline{u}$ utters $\underline{h}$' then we have simply refused to play according to the rules of the game which require that $\underline{J}$ understand the fore-diction in question. We shall later note that there is substance to the objection. But before examining this we need to take note of another difficulty.

We remember that $\underline{u}$'s problem is to successfully predict $\underline{J}$'s behavior. 6.21 and 6.11 do not, by themselves, enable $\underline{u}$ to predict what $\underline{J}$ will actually do. What is required is that the antecedent circumstances also be specified. But in these two cases what is required is that $\underline{u}$ actually utter a statement which is false, or more specifically a self-vitiating hypothesis, for the antecedent of 6.11 or 6.21 to be fulfilled. It might therefore be suspected that statements from which other statements such as 6.11 and 6.21 were deducible would be unsuitable. We need only note, however, that any theory of self-vitiation would be required to enable us to deduce statements of the form 6.11 and 6.21 if the theory would be adequate in the sense of enabling us to make a preponderance of successful predictions. What we shall

attempt to show is that the difficulty which we encounter in such cases as that of J is an empirical and not a logical one.[5]

The solution we propose depends on our lifting the kinds of restrictions which we have purposely imposed on the prediction of J's behavior by u in order to bring out the nature of the problem involved. What our example has shown is not that J's behavior is unpredictable or that there is a logical problem which is unavoidable in such cases. The difficulty lies in making the requirements for a successful prediction so rigid as to introduce the very perturbance which creates the difficulty. In this regard our Mars example may prove illuminating. Suppose, for example, that we required that the astrophysicist in making predictions of the behavior of Mars were to utter his fore-dictions only in the form of explosions of cobalt bombs as described in Chapter II. This absurd requirement would mean that very few, if any,

---

[5] We may admit that at present this kind of difficulty has little practical import. It may not be one which will occur with any frequency. Yet it is precisely such "theoretical" difficulties which when resolved help to increase our understanding.

On the other hand, it is by no means certain that a theory of self-vitiation, if used with any frequency, would not tend to increase the occurrence of such cases as that of J. It may even be suspected that such a theory if used in certain areas of human concern would tend to generate an attitude of defiance on the part of those whose behavior was being predicted.

of the predictions concerning Mars' behavior would turn out to be successful. Now the requirement that predictions of the behavior of individuals like J be known and understood by the individuals in question seems as absurd, from a logical or epistemological standpoint as that of requiring that predictions of behavior of Mars contain only those fore-dictions which explode in the fashion described earlier.

We are not claiming that the requirement is absurd per se. The difference lies in the fact that human beings, unlike planets, dislike being used as guinea pigs. And it may well be an empirical fact that under certain circumstances the goal of attaining a successful prediction may need to give way to other goals with which it is empirically incompatible. We do not wish to dispute or to deny this possibility. As we indicated earlier, we are not so much concerned with the practical as with the theoretical aspects of the problem of self-vitiation. Any attempt to do justice to the practical aspects would perhaps require another treatise. Without wishing to deny that the cost of successful prediction may be very high in terms of human values depending on other circumstances,[6] we are confining our attention to the question as to whether there are any logical difficulties to predicting human behavior.

---

[6]See Merton's comments quoted on p. 56.

Suppose now that we have an unseen observer O who watches the drama between u and J. In the event that O has evidence of the assertion that u utters h(sp) we note that he can predict (and presumably successfully predict) J's subsequent behavior. He does so on the basis of 6.11 and the fact that u has uttered h(sp). Similar remarks would hold if u uttered $\tilde{h}$(sp).

Involved here is the possibility of a sequence which is very like that of the one elaborated concerning fatalities on Hol. If O makes his prediction known to J, J may choose to vitiate O's prediction; but in so doing he will confirm u's prediction insofar as the two fore-dictions in question are contradictories. Likewise R may predict that O's fore-diction is self-vitiating; and if this latter fore-diction is not known to J, it may be presumed to be more acceptable than that of the other two inasmuch as it considers boundary conditions not considered by u and O, namely the fact of their utterances including J as a receiver.

Again we note that self-vitiating phenomena do not give us grounds for asserting that any event is unpredictable in principle. Whether or not we can successfully predict the behavior of any individual or group of individuals on the basis of a theory T will depend on whether or not we are able to include the relevant boundary conditions in the explanans of the prediction.

Whether we choose to accept O's fore-diction or u's fore-diction will depend primarily on which of the two fore-dictions we consider to be better evidenced. Again what is required is a theory of confirmation which will enable us to answer this question.

An Empirical Requirement for Successful Prediction

The term 'successful prediction' is subject to all of the difficulties encountered in those terms which share a process-product ambiguity. 'Prediction' may be understood as referring to the process whereby a particular scientist or group of scientists deduce certain conclusions from a set of law-statements and antecedent condition statements. It may also refer to the product which is the outcome of this process: i.e., the set of tokens thus produced. The term 'successful prediction' likewise demands clarification. A prediction may be considered successful if the conclusion actually follows from the premises and unsuccessful if it does not follow from the premises or if the process is interrupted and the conclusion never derived. It is not, of course, this sense of the term with which we are primarily concerned. The way in which we have been employing the term is as follows: A prediction P is successful only if the fore-diction h of P is true.

The empirical requirement for a successful prediction with which we are concerned may be stated as follows:

6.5 A prediction P is successful only if it does not contain a self-vitiating fore-diction.

6.5 stipulates a necessary (but not a sufficient) condition for a prediction being successful in the sense explicated. If a prediction is successful we know that it does not contain a self-vitiating fore-diction. But the fact that a prediction does not contain a self-vitiating fore-diction is not sufficient grounds for believing it is successful. Predictions fail for reasons other than that of containing self-vitiating fore-dictions. (For example, they may also fail because they contain a self-vitiating premise.)

By this stage in our analysis the requirement for successful prediction may appear so obvious as to be trivial. But should it appear trivial we suggest that its apparent triviality is due in part to our having been able to shed a little light on the problems of self-vitiation.

To return to u and his obstreperous receiver we can see that the problem is precisely that of attempting to make a successful prediction while ignoring requirement 6.5. What is being demanded of u is that which is presumably not empirically (and in one sense, logically) possible -- namely, that u successfully predict J's behavior using a self-vitiating fore-diction. As we pointed out earlier, such a demand is as unreasonable as a demand that astrophysicists make all the utterances of

of their fore-dictions in a manner which would upset the causal sequence anticipated by the fore-diction.

We have suggested that u would not predict both h and $\tilde{h}$ on the basis of the same theory since this would demonstrate that theory to be inadequate. But suppose that u unmindful of certain logical niceties decides that the outcome of his prediction is not particularly important in the "total scheme of things" and acts as follows: he utters h(sp) which he has deduced from a body of psychological theory. But then on the basis of the theory of self-vitiation T' and the fact that he has already uttered h(sp) (i.e. on the basis of 6.11) he utters $\tilde{h}$(sp). Thus far he has uttered a contradiction. But he has done so on the basis of two different theories. He knows therefore that there is something wrong with at least one of the theories in question. Just which theory is to be accepted will depend on which of the two theories in question is better-evidenced. And this problem as we noted previously is one which at present has no acceptable solution. But now on the basis of 6.21 plus the fact that he has uttered $\tilde{h}$(sp) he can also deduce h_1(sp) where h_1(sp) is a replica of h(sp). But now he has deduced a contradiction on the basis of just one and the same theory, T'. Hence 6.11 and 6.21 are inadequate as they stand. One possible solution to the problem is as follows:

6.13 [6.11 . 6.21] if and only if u does not utter both h and $\tilde{h}$.

That is to say, we can assert the conjunction of 6.11 and 6.21 if and only if u utters only one of a pair of contradictory sentences. This new formulation 6.13 does at least two important things. First, it prevents the kind of contradictions which we have seen could be deduced from T' if 6.11 and 6.21 were allowed to stand unqualified. Second it preserves 6.11 and 6.21 for the theory. As we have seen, any adequate theory of self-vitiation would need to include 6.11 and 6.21.

But while 6.13 enables us to preserve the theory it does not enable u to predict the behavior of J. u is "right back where he started." He knows that either utterance he makes will be self-vitiating. At this point he may give up in despair, strike J in frustration, or utter contradictions until he is tired of the game. 6.5, however, elucidates his problem. One cannot utter a fore-diction of a successful prediction when the utterance of such a fore-diction is self-vitiating. This is not to say that u cannot successfully predict the behavior of J. It is only to say that he can successfully predict the behavior of J when J does not understand or misunderstands or is not aware of what u has predicted.[7]

---

[7]A number of interesting variations on this theme are possible. Suppose that u utters h. But then to himself and unbeknownst to J he deduces $\tilde{h}$ on the basis of 6.11 and the fact that he has uttered h but not $\tilde{h}$. Now suppose that J is aware of the laws of self-vitiation and is also aware of 6.5 as a requirement. He will now be puzzled as to how to act. If he acts as h describes, he has failed to make good his claim. But if he acts as $\tilde{h}$ describes, then he is prepared to

We have not discovered anything new in 6.5. Venn was perfectly aware of this solution when he said:

> The publication of any prediction about the conduct of human beings (unless it were kept out of their sight, or expressed in unintelligible language) almost certainly would have some effect.[8]

Venn's comment is made parenthetically and it is not clear as to whether he would be willing to allow that certain scientific predictions ought to be expressed in unintelligible language or kept out of the sight of those about whom the prediction was made. The point is similar to our own: Successful predictions do not contain self-vitiating fore-dictions. We do not, however, draw Venn's conclusion. He believed that this problem made it difficult (impossible?) to predict the behavior of human beings.[9] He also believed that the phenomena in question enabled us to distinguish the social from the physical sciences. We can agree that self-vitiation is a difficulty in the social sciences. But we believe it is a difficulty to be overcome, not one to be tolerated.

---

have u exhibit 6.11 and 6.5 as the basis for his "real" prediction. The puzzler has now become the puzzled.

Analogous problems are dealt with in the theory of games.

[8] Op. cit., p. 576

[9] Ibid.

Merton, while apparently more concerned with value questions, has also pointed out the solution we have indicated in 6.5.[10] But Merton is perhaps so concerned with the value questions that he has failed to note some of the implications of his position. Some of these implications may be suggested by way of an example.

Suppose that a social scientist K after investigating racial tensions in the United States up to 1962 concludes that all the major problems of race relations will be resolved by 1970. As a scientist he wishes to publish his findings. Scientist F is more cautious. He notes that the fore-diction of K has characteristics A, B and D, i.e. it is a self-vitiatable utterance. And should circumstances C obtain, the utterance will be self-vitiating. He urges K not to publish his results inasmuch as the mere publication will cause the racial tensions to be prolonged. K, however, as a scientist, claims that he has an obligation to make known what he believes to be the truth. F, however, points out that the fore-diction of K is not the truth. What is the truth is better expressed as, "Either racial tensions will be minimized by 1970 or the utterance of K in 1962 regarding the minimizing of racial tensions in 1970 is self-vitiating if K publishes replicas of the sentential

---

[10]See footnote, p. 56, this essay.

parts of his utterance." In other words, K's utterance is true only if certain antecedent conditions obtain. And the most important, presumably, in this example, is that replicas of the sentential parts of his utterance not be published.[11]

---

[11] K's utterance contains as sentential part a replica of:

6.7 All the major problems of race relations will be resolved by 1970.

Suppose that F were then to publish:

6.8 K's utterance is true only if no replicas of it are published prior to 1970.

Whether or not 6.8 is self-vitiating would depend on the laws of self-vitiation and other antecedent circumstances. The answer to this question would depend in part on the form of F's utterance. Does it characterize, name or state what is contained in K's utterance? Certain semantic problems may be involved also. Suppose F publishes his assertion in the following form:

6.81 K's utterance that all the major problems of race relations will be resolved by 1970 is true only if no replicas of it are published prior to 1970.

Does 6.81 constitute publication of K's utterance? What would be needed in order to answer this question is a clarification of what is included in the extension of the term 'publication'. A possible difficulty could be that of the self-referential feature of 6.81. It speaks of the publication of replicas of 6.7 while apparently containing a replica of 6.7. This self-referential feature could present interesting problems in itself. But we have already shown how it is always possible to avoid this kind of difficulty by using names of tokens (or utterances) rather than replicas of them. (See our discussion of this issue in Chapter II.) And presumably a scientist more fastidious in his choice of tokens than F would seek to avoid the additional problems created by making fore-dictions in the form of 6.81.

The question of the self-vitiating character of 6.8, however, would be in the final analysis an empirical one.

For K to publish his utterance would then be irresponsible on two counts: he would be publishing that which he had empirical reasons for believing was a false utterance (i.e. after F had confronted him with the theory of self-vitiation and the facts) and he would be contributing to the increase rather than the decrease of racial tensions. In neither case would the publication of the results appear to be justified.

This analysis would suggest the rather far reaching conclusion that the interests of science are not always furthered by the publication of scientific results. Nor are the human values we mean to preserve in our day to day affairs always preserved by knowing what purports to be the truth. The implications for value-theory cannot be overestimated. In one sense the scientist appears to be in the middle of a dilemma. He knows the truth. But if he publishes what he knows as the truth, the antecedent conditions regarding boundary conditions are changed and it is no longer a true but a false sentence which he has uttered.

Apparently, until an adequate theory of self-vitiation is forthcoming, each scientist must make his decisions on the basis of his own intuition and what empirical evidence is available in regard to self-vitiating utterances. The contemprary social scientist is in somewhat of a predicament. He may suspect that some of his utterances will have self-vitiating consequences

under conditions of publication. Yet the present set of values by which the scientific community is apparently guided would lead him to publish his results provided that he had good empirical evidence for his conclusions.

But while the phenomena of self-vitiation puts the scientist in a predicament, the phenomena of self-fulfilling hypotheses may offer some modicum of reprieve from the tensions engendered by the phenomena of self-vitiation. Since we at present do not have a theory of self-vitiation, we cannot do more than suspect or guess or prophesy that a particular utterance will be self-vitiating. It may turn out fortuitously, however, that some of the utterances of the social scientist will tend to be self-fulfilling and that unless F can produce an adequate theory for his assertions concerning K's utterance, K may, in the absence of empirical evidence, publish his results with no more than a normal amount of compunction attendant upon the publication of an empirical hypothesis.

We are not suggesting that the problem is therefore unimportant. Our whole essay has been concerned to point up the great importance of this phenomenon for science and derivatively for humanity at large. Our point is simply that there is as yet no empirical justification for keeping a large number of hypotheses secret. We do not know, on the one hand, which utterances are self-fulfilling and which self-vitiating and on the other

hand, which have relatively little causal significance or import. And until an adequate theory is available, we are perhaps best guided by those values which have guided the scientific community in the past. In the event that an individual scientist suspects that his hypothesis would be self-vitiating if uttered, then he shall have to weigh the possible costs involved in publishing or in refusing to publish. Perhaps this is one price of being human.

Self-vitiating Instances of a Theory of Self-vitiation

An interesting problem arises out of the preceding discussion. We can see that a theory of self-vitiation itself is not necessarily immune to the problems considered in this essay. We have had occasion to note that predictions concerning the self-vitiating character of specific utterances may themselves be self-vitiating. And if the same theory of self-vitiation is used in predictions which occur on two successive levels, we recall that one or the other of the two fore-dictions is false. We have observed that in such an instance what will usually be the case is that a "higher-level" prediction will usually take account of certain relevant boundary conditions not considered by the "lower-level" prediction. But any time such a set of circumstances obtain the

theory in question will be considered inadequate since it will be possible to deduce a contradiction from the theory.[12] Since this difficulty presumably would be one which any theory of self-vitiation would have to deal with we may inquire as to whether or not there are any general principles which can be applied to a theory of self-vitiation which will prevent the kind of difficulty here encountered. An obvious solution is suggested by 6.5 which constitutes an empirical requirement for successful fore-dictions. This requirement can be extended to cover the generalizations of a theory as follows:

$$6.6 \quad (x)(x \epsilon H \supset Tx) \supset -(\exists x)(x \epsilon H \,.\, Vx)$$

Requirement 6.6 constitutes a necessary condition for the truth of a generalization. It does not constitute a sufficient condition for it does not take account of all the possible boundary conditions which could be considered as disconfirming evidence of the truth of H.

Requirement 6.6 is useful in any theory. We are prepared to assert the truth of a generalization only if it contains no member which is self-vitiating. But 6.6 is especially important to a theory of self-vitiation inasmuch as we may expect that such a theory would tend to generate predictions on several levels.

---

[12] See pp. 169-171 for an outline of a proof.

One implication of this requirement is that for the purposes of deduction no generalization in a theory of self-vitiation would be asserted without the qualification of 6.6. This statement may appear to be trivial inasmuch as an $\underline{H}$ containing a self-vitiating member would not be considered to be true.[13] But there appears to be a self-referential puzzle here involved which is by no means trivial. Consider the following circumstances: a scientist wishes to know the truth of a generalization $\underline{H}_g$ in a theory of self-vitiation $\underline{T}'$. One of the conditions for the truth of $\underline{H}_g$ would be that it not have any self-vitiating members. How shall he proceed to determine whether or not $\underline{H}_g$ has any self-vitiating members? He shall have to use theory $\underline{T}'$ in order to determine this. But in making this determination he may have to rely on the truth of $\underline{H}_g$, due to the fact that $\underline{H}_g$ is part of $\underline{T}'$. He is thus involved in a circular quest after truth.

An immediately apparent solution to this problem is one we have encountered earlier. What is now needed is a theory $\underline{T}''$ by means of which one can determine the truth of $\underline{H}_g$. $\underline{H}_g$ we assume is not part of $\underline{T}''$ although it is part of $\underline{T}'$. Hence both the self-referential features of the problem are avoided as well as the circular nature of the quest.

---

[13] These statements are of course subject to the qualifications made earlier regarding the cost of abandoning a generalization when it has proved exceedingly useful in the past.

And should it be necessary to determine the truth of a specific law-statement of T'' we would then be required to make this determination on the basis of a theory T'''. Just how far we would be prepared to go in developing theories which explain theories which ... explain theories would depend in large measure on how important for our purposes we considered the enterprise to be. And, as we indicated earlier in connection with predictions on various levels, there may be certain psychological limitations which would not make it required that we go much beyond T''' in determining the truth or falsity of the generalizations whereby we make our predictions concerning self-vitiating phenomena. And we may suspect, presystematically, that generalizations of T''' would have few if any self-vitiating instances.[14]

Again we note that in our analysis we have discovered no theoretical obstacles to the development of an adequate empirical theory of self-vitiation.

---

[14] Another solution to this particular problem would be to develop a general theory of the theories of self-vitiation. But this general theory would be subject to the same kinds of questions raised above. Perhaps, however, such a general theory would need to proceed to fewer levels to achieve adequacy than does the schema presented above. But this argument has, we may fear, gone beyond practical significance. At present we would be grateful for a theory such as T'. Once this has been highly confirmed we may then undertake the construction of T'' and then perhaps a general theory which explains certain features of both. Until we possess T' much of what we are considering in this chapter is too speculative to be of much value. We are only concerned to show that these difficulties are not theoretically insurmountable.

Predicting Our Own Behavior

One final problem needs to be examined; it involves the possibility of predicting our own behavior. We have ignored this general area until now because previously our analysis was incapable of dealing with the problem. The problem can be phrased in the form of a question: "Can I, as a being who is capable of making certain voluntary choices, predict how I will act under certain conditions and circumstances?" Henle claimed that we cannot predict our own behavior because such a prediction is tantamount to a decision.[15] Richfield and Copi, in replying to Henle, have argued that decisions and predictions of our own behavior represent two distinct classes of phenomena. The distinction in question is clarified by an example:

> Suppose I decide to eat lunch early tomorrow, for one reason or another. Then I remember that tomorrow morning is scheduled be very busy [sic] , and I predict that the press of work will be such as to make me decide then to delay my lunch to a later time. Still I hope that the work can be got through early, and I renew my decision to eat lunch early rather than late. Here we have both decisions and prediction with respect to one's own future voluntary action. And one is not "really" the other because they are, in this case contraries.[16]

On the other hand there are circumstances under which we discover that decision and prediction need to be compatible.

---

[15]Paul Henle, Op. cit.

[16]Op. cit., p. 48.

> My son would like bubble gum and a purple pony. I wish to please him. I promise him the gum, but I cannot promise him a purple pony. As an upright father, I cannot decide to act in a way that I cannot also predict. If, as Professor Henle suggests, what I decide I cannot also predict, then it is hard to see the basis upon which I could ever grant my assurances of my own actions.[17]

Now the interesting cases for our purposes are those in which predictions of our own behavior turn out to be self-vitiating. I, for example, am one who is adicted to nicotine and smoke two packs of cigarettes each day. On the basis of evidence and certain laws of human behavior I can quite accurately predict that tomorrow I will smoke two packs of cigarettes. On the other hand, I can also decide to quit smoking tomorrow. That is, I can choose to vitiate my prediction. The question then would be whether or not I could predict how I will decide. Will I decide to stop smoking and hence to vitiate my original prediction? Or will I simply follow the course of least resistance? In order to answer this question I would need a theory of self-vitiation which would be able to determine whether my original forediction was self-vitiating, self-fulfilling, or neither.[18] And if I am extremely anxious to predict how I will act,

---

[17] Ibid.

[18] We are assuming that an adequate theory of self-vitiation would also be able to account for the phenomena which we have labeled 'self-fulfillable'.

I may also wish to inquire whether this second fore-diction is itself self-vitiating. I may then need to resort to theory T'' assuming that theory T' was used in the second prediction. It may also be the case that I will make certain decisions on the basis of what I have predicted from theory T'', i.e., I may choose to vitiate what I have predicted or I may follow the line of least resistance.

At this point an ancient philosophical controversy comes to the fore: the controversy regarding the primacy of the freedom of the will versus the primacy of reason. And whether I choose to vitiate a fore-diction on any "level" would presumably be a question of whether I were more interested in knowing what the future would bring or whether I were more interested in being able to decide what the future would bring. And we may suppose that an adequate theory of self-vitiation would, on the basis of which of these dispositions was manifested by a particular receiver of tokens, be able to determine on which level an individual would stop deciding and predicting.

From the analysis of self-vitiation we can see that whether or not predictions of our own behavior are successful will depend in large part on whether or not the fore-dictions in question are self-vitiating. It is certainly true that I make a number of successful predictions about my own behavior every day. These are not

simply decisions in which I choose one alternative over another. I deduce on the basis of certain inductive generalizations concerning my own and others' behavior that I will act in such and such a way under certain conditions. I predict that I will go to work tomorrow, that I will eat at least three meals, that I will exhibit certain moods, that if I hear a joke I will laugh, that if I hear of the death of a friend I will be sad, that I will continue to smoke two packs a day, that in the next five-years I will probably spend at least five days in a sick bed, that I probably will not die tomorrow, that tomorrow I will perform certain bodily functions, etc., etc.

This is not to say that in all of these events I may not choose to do otherwise. But in the case that I do choose to do otherwise I can then raise the question as to whether I can successfully predict how I will choose. And it may turn out that certain kinds of personal predictions are unsuccessful and <u>not empirically possible</u> due to the fact that they introduce the kinds of disturbance with which we are now quite familiar. In this case we cannot retreat to an unseen observer in order to resolve the problem. For here we are dealing with predictions of one's own behavior. But which of such predictions are or would be unsuccessful is itself an empirical question which could be answered by an adequate theory of self-vitiation. Again, we note however, that

this is not to say that any action is unpredictable in principle; only that when certain predictors are involved may the action prove to be unpredictable.

But even in such cases the individual predicting his own behavior may choose one level of prediction over another on empirical grounds. And just because a prediction on one level contained a self-vitiating fore-diction would not be adequate grounds for claiming that fore-dictions on every level would be self-vitiating.

Again, 6.5 requires that successful predictions must not contain self-vitiating fore-dictions.[19] And whether an individual can successfully predict his own behavior in any set of specific circumstances will depend to a large extent on whether his utterances of fore-dictions are self-vitiating.

Summary

The sections of this chapter have been concerned with arguments previously considered in Chapter II. We have been principally concerned to show that the arguments claiming that the phenomena of self-vitiation makes successful prediction of the future behavior of human beings impossible are not cogent. The difficulties involved in the development of a theory of self-vitiation

---

[19]Although obviously they may contain self-fulfilling fore-dictions.

are formidable indeed. But the difficulties involved are not logical but empirical.

We went on to show that there are certain empirical conditions which must be met if a particular prediction is to be accepted as successful. The result of this analysis has been to show that the success or failure of a prediction may depend on the circumstances under which a particular fore-diction is uttered (e.g. the relations which obtain between the predictor and the receiver of tokens).

The purposes for which this analysis was undertaken have, we hope, been fulfilled. Whatever differences do obtain between the social and physical sciences, it seems clear that the phenomena of self-vitiation is not one of them. It also seems clear that there are no theoretical or logical obstacles in the way of developing a theory of self-vitiation. Or, more modestly, we may say that if there are logical difficulties involved, we have been unable to uncover them.

CHAPTER VII

SUMMARY AND IMPLICATIONS

Summary of the Analysis

In this essay we have attempted to call attention to the great importance of the phenomena of self-vitiation for the philosophy of science. What we can know about the world is in some degree dependent on which of our pretensions to knowledge have self-vitiating consequences. In this chapter we shall review our argument and then attempt to call attention, in somewhat more detail, to some of the consequences of what we have presumably learned from the analysis.

We first examined a number of instances of self-vitiation. While these examples were highly selective, they suggest that there appears to be little of human concern which is immune to the possibility of being described by self-vitiating utterances.

The arguments in the literature concerning the impossibility of predicting the future history of mankind we found to be incogent for a number of reasons. The phenomena with which the arguments were concerned were not clearly specified. Certain ambiguities in key terms prevented the arguments from having the cogency usually attributed to them.

These difficulties led us to undertake a more formal and more rigorously precise analysis of the term 'self-vitiating'. We defined the term in such a manner that the definiens enabled us to discriminate between

those utterances which are simply false and those which we are warranted in accepting as true prior to the determination of their self-vitiating character.

We also called attention to the fact that this class of utterances needs to be distinguished from the traditional syntactical and semantical paradoxes. It is no self-referential feature of the self-vitiating utterance which creates the difficulty. Nor is it a syntactical feature of these utterances which proves troublesome. What is involved is that the fore-diction in question is causally efficacious in bringing about a state of affairs described of the negation of the utterance. That this distinction has not been clearly drawn in the past may account for the rather obscure fashion in which the phenomena have been treated.

We found that the difficulties encountered with certain malformed (from a syntactical standpoint) and ambiguous or vague expressions could be overcome by treating such utterances as members of the causal sequence which leads to a self-vitiating consequence and reserving the predicate 'self-vitiating' for well-formed, unambiguous and precise expressions. This restricted usage, we assumed, will not handicap the social scientist in developing a theory of self-vitiation.

One of the consequences of the analysis is that usually only certain members of a statement are self-vitiating. If we say that a statement is self-vitiating,

this loose manner of speaking may be interpreted as meaning that some particular member of that statement is self-vitiating but not that every member is self-vitiating. Hence if some member of H is false, every member of H is false. But it is not the case that if some member of H is self-vitiating, every member is self-vitiating.

We were fortunate in discovering a mode of analysis for the predicate 'self-vitiating' which also proved to be adequate to the analysis of the predicate 'self-fulfilling'. The analysis, however, revealed an important distinction between the two predicates. This difference concerns the predicates as applied to generalizations. We noted that, from a logical standpoint, one disconfirming instance of a generalization which was caused by an utterance of that generalization would be sufficient for asserting that the generalization is false. But one confirming instance of a contingent generalization which was caused by the utterance of that generalization would not be sufficient grounds for holding that the generalization is true.

Our analysis also showed certain important relations which exist between the phenomena of self-vitiation and the problem of distinguishing between law-like and non-law-like generalizations. We were led to suggest that an adequate treatment of the one would contribute to a solution of the other. If the grounds

which are sufficient for determining that e occurs are also sufficient for determining the truth of h, then in the event that h is a generalization we could conclude that h is non-law-like.

An examination of the problem of boundary conditions in scientific explanation revealed that self-vitiating phenomena may be treated as a sub-class of that set of problems encountered when the boundary conditions are not made explicit in a particular explanation. The difficulty in these cases is removed by making the boundary conditions explicit. From the revised set of premises the false explanandum is no longer deducible.

An adequate theory of self-vitiation would explain why certain predictions fail; i.e. it would show that one of the boundary conditions not considered in such cases was the self-vitiating character of the fore-diction. Even more important, perhaps, would be the fact that an adequate theory of self-vitiation would enable us to predict the self-vitiating consequences of certain utterances. As part of a general theory of perturbance such a theory of self-vitiation would be invaluable to both the social and the physical sciences.

Prior to the development of a theory of self-vitiation it would be required to examine a large number of instances of self-vitiating hypotheses and then to determine which features shared in common by all of these utterances are relevant for defining the disposition

predicate 'self-vitiatable'. Once such a predicate is available, it will be possible to begin formulating the laws of self-vitiation, gathering evidence for the confirmation of the law-statements in question, and then systematically ordering these statements.

The practical difficulties attending the development of such a theory would be enormous. We have attempted to show, however, that the arguments concerning the supposed logical difficulties attending the development of such a theory will not stand up under examination. The lack of an adequate theory at present is a major factor in our being unable to successfully predict the future course of human history. But there are no apparent theoretical reasons why such a theory could not be developed. And if such theoretical reasons "exist" such reasons have not yet been advanced in favor of the thesis of the impossibility of predicting certain events. The arguments which <u>have been advanced</u> do not adequately support this thesis.

Our analysis has also attempted to show that the claim that the phenomena of self-vitiation enable us to distinguish between the social and non-social sciences is not as cogent as it may first appear. Any difference which does emerge seems to be one of degree rather than of kind. That is to say, we may expect, presystematically at least, that we will discover more examples of this type of phenomenon in the social sciences than in the

physical sciences. We noted, however, that as greater use is made of machines such as electronic brains the difficulty may become even more important for the physical or non-social sciences as well.

In a positive vein our analysis attempted to indicate some of the problems that the development of such a theory could presumably resolve. Without an adequate theory of self-vitiation certain advances in the sciences do not appear to have much chance for success. With such a theory, however, we may expect that our understanding of the world in which we live will be greatly increased.

Some Implications of the Preceding Analysis

We have suggested throughout the course of this essay that the phenomena of self-vitiation is not simply an interesting though somewhat irrelevant problem but that it has great importance for the sciences and the enterprise of human living generally. We shall in this concluding section attempt to indicate in somewhat more detail some of the practical implications of the development of such a theory.

A theory of self-vitiation would presumably contribute to a more adequate theory of knowledge. In Chapter II we showed some of the difficulties attendant upon our usage of the term 'to know'. (We did not pretend to give an exhaustive account of the problem in that section, nor can we attempt such an account here.

We wish merely to indicate some of the other consequences of our analysis.) When we say that we know that such-and-such is the case, if an utterance of a statement which describes what we claim to know is self-vitiatable, how can our claim to know such-and-such be construed? The preceding analysis indicates that such claims may be elliptical ways of saying that such-and-such is the case, provided that so-and-so is not the case. In other words, all absolute claims to knowledge (with the possible exception of such claims as "I know that the paper before me appears to be white") may be construed as provisional claims. What is tacitly assumed is that certain boundary conditions will hold. The interesting cases for a theory of knowledge are just those cases in which the utterance of what is construed to be knowledge is self-vitiating. One boundary condition which is presumed to hold in this case would be that the utterance not occur under circumstances C.

Presumably the fore-going remarks will only apply to contingent or synthetic hypotheses and not to those which are logically true or analytic. But the kinds of distinctions which are required here are by no means clear. The literature of the past few years is filled with controversy concerning the distinction between formal and non-formal, analytic and synthetic, contingent and necessary statements. Our analysis may or may not help to settle this controversy. But some of the

difficulty may be resolved by noting that should any statement contain a self-vitiatable member we should presumably not wish to count that statement as analytic. And if a statement contain no self-vitiatable member we should probably not wish to call that statement synthetic.[1] Here we would have an empirical test available for determining whether to apply the predicate 'analytic' in any given case. An analytic statement then would be one which had no self-vitiatable members -- i.e. no members such that if they occurred under circumstances C would be self-vitiating. An interesting consequence of this treatment would be that to call a particular statement 'analytic' would be to claim that by law it was not possible for the statement to be false. The possibility here in question would be "empirical" as well as "logical" or "mathematical". The laws in question would be the laws of self-vitiation together with all the other laws of the sciences. We are not attempting to defend the distinction between analytic and synthetic as it has traditionally been drawn. We are only attempting to show that a theory of self-vitiation could presumably enable one to mark the distinction on empirical grounds. Whether the distinction as thus drawn would prove useful is not a question we need to decide in this essay.

---

[1] Except perhaps in certain cases such as statements in celestial mechanics of the kind indicated in Chapter V.

The implications of our analysis for each of the sciences cannot, it appears, be overestimated. One of the requirements which the scientist imposes upon his enterprise is that disturbing factors, perturbance effects, and relevant differences between the subjects under consideration be reduced to a minimum. But is there any way open to the scientist to prevent the disturbance which is the result of simply drawing a conclusion (i.e. uttering a self-vitiating fore-diction) from a set of premises describing the results of his experiment? A theory of self-vitiation presumably could answer this question for the individual scientist in specific circumstances. What would be available from such a theory would be a set of requirements for uttering fore-dictions which would tend to minimize the possibility of their being self-vitiating. We have already suggested what at least one of these requirements might be: a keeping of the results as secret as possible, preventing them from being understood by receivers who have the disposition to render them self-vitiating.

And here we note some important value problems which need to be considered. One problem is at least as old as Plato's Republic and the question of whether or not the beneficial lie is justified and the conditions under which it would be justified. In a democratic society the decision to keep certain results secret appears to violate some of the values which a free society

holds dear. We may even equate the kind of secrecy in question with totalitarian and authoritarian regimes. Once a theory of self-vitiation is available which can enable us to predict successfully in a large number of cases when certain utterances are or are not self-vitiating, then as we noted in Chapter VI, part of the value problems are solved. For if a scientist keeps secret a conclusion which he knows will be self-vitiating if uttered under certain circumstances, circumstances which would be almost certain to obtain unless he kept the conclusion secret, then he is not hiding the truth, but refusing to publish what will in those circumstances be false.

The major value problems are present ones. We do not have a theory of self-vitiation; even an inadequate theory is not available to us. Hence we are unable in any given set of circumstances to determine with any significant degree of assurance when a particular utterance will be self-vitiating. Yet we may strongly suspect, guess, or prophesy that certain utterances will have self-vitiating consequences. What shall we do in such cases? The answer to this question is by no means clear cut. And until an adequate theory is forthcoming we shall probably have to rely on guesses concerning probabilities, weighing the costs of uttering over against the cost of not uttering certain hypotheses.

The political and social consequences appear to be enormous. If a political scientist or a sociologist publishes certain findings concerning race relations, unemployment, labor-management relations or any other of a number of related issues, and these turn out to have self-vitiating consequences which are disvaluable according to some acceptable value schema, the entire society of which the scientist is a part will presumably suffer the consequences. How often such publications of this kind have such effects is at present unknown to us. What is clear, however, is that the individual scientist's dedication to the truth is not always exhibited by his publishing what purports to be the truth.

On the other hand suppose a social scientist deduces a conclusion which if published would be self-vitiating yet among the consequences of publication would also be certain valuable results (e.g. the amelioration of race relations, the decrease in unemployment, etc.). Suppose further that the scientist has evidence that his hypothesis is self-vitiatable and would be self-vitiating if published. Shall he publish a lie? Plato's question has not yet been satisfactorily answered; and presumably it cannot be given a blanket answer.

The present day scientist, however, does not have the kind of evidence which we would require to support an assertion that a particular scientific hypothesis is self-vitiating or is not self-vitiating. And whether a given

scientist will publish results of his particular investigations may in part depend on how concerned he is about the possibility that his publication of results may have self-vitiating consequences.

An even more troublesome problem arises in regard to confirmation theory. It is likely that a theory of self-vitiation will give us a set of probability statements concerning the phenomena of self-vitiation. Suppose that a scientist, on the basis of such a theory determines that a particular conclusion, if published, has a probability of .7 of being self-vitiating. Shall he publish the fore-diction? Or suppose that instead the probability in question is .15. Shall he publish the fore-diction? Whether or not he will publish may depend on the consequences of publication. If the fore-diction in question has consequences upon publication which are relatively unimportant in the "total scheme of things" a probability of .7 may not dissuade him from publication. But a fore-diction whose publication may tend to result in atomic war if it is self-vitiating, may have a probability as low as .01 of being self-vitiating and yet the scientist may still be hesitant to publish.[2] Publishers also may be called

---

[2]These comments are suggested by and are related to Professor Richard S. Rudner's comments in "The Scientist Qua Scientist Makes Value Judgments", Philosophy of Science, 20:1-6, January, 1953. One of Professor Rudner's main points is that "how sure we must be before we accept a hypothesis depends on how serious a mistake would be."

upon to make certain value judgments in this general area. Interesting questions of a legal nature may arise once an adequate theory of self-vitiation is available. Suppose that the publication of a particular hypothesis had a presumed probability of .7 that it was self-vitiating. Suppose further that the author and publisher know (in one of the pertinent senses discussed earlier) that if the hypothesis is self-vitiating it will have consequences which are extremely disvaluable to a certain segment of the population. If author and publisher decide to go through with publication will they then be liable to lawsuit for damages? If not, would the same hold true if the probability were .99? If so, would the same hold true if the probability were .01?

These remarks indicate that certain delicate problems of censorship may be given a slightly different "twist" once an adequate theory of self-vitiation is available to us. We may suspect that certain legal restrictions on publications will be urged by those who emphasize the consequences of certain kinds of publication. Others may hold that the decision to publish or not to publish must rest with the individual author and publisher and that the freedom of the press must not suffer any more inroads than it does at present. How this particular argument will be resolved is by no means clear. We are concerned at this point only to call attention to one

possible consequence which a theory of self-vitiation may have once it has proven adequate for other purposes.

The problems which an adequate theory of self-vitiation will presumably pose are by no means new. And we have not pretended to do more than to suggest a few of what appear to be some of the major difficulties we shall encounter once such a theory is available to us.

Our task of analyzing the terms 'self-vitiating', 'self-vitiatable' and 'self-fulfilling' is for our purposes complete. The task of developing an adequate theory of self-vitiation remains. And if this essay should prove to be a help in the development of such a theory our hopes shall have been fulfilled.

BIBLIOGRAPHY

Books

Berlin, Isaiah. Historical Inevitability. London and New York: Oxford University Press, 1954.

Braithwaite, Richard Bevan. Scientific Explanation. New York: Harper and Brothers, 1960.

Carroll, Lewis. Through the Looking Glass. New York: Random House, special edition, 1946.

Danto, Arthur and Sidney Morgenbesser, editors. Philosophy of Science. New York: Meridian Books, Inc., 1960.

Dray, William. Laws and Explanation in History. London: Oxford University Press, 1957.

Goodman, Nelson. Fact, Fiction and Forecast. Cambridge, Massachusetts: Harvard University Press, 1955.

_____. The Structure of Appearance. Cambridge, Massachusetts: Harvard University Press, 1951.

Leonard, Henry S. Principles of Right Reason. New York: Henry Holt and Company, c.d., 1957.

_____. Singular Terms. typescript, Widener Library, 1930.

Luce, R. D. and H. Raiffa.. Games and Decisions. New York: Wiley and Sons, Inc., 1957.

McKinsey, John C. C. Introduction to the Theory of Games. New York: McGraw-Hill Book Company, Inc., 1952.

Madden, Edward W. The Structure of Scientific Thought. Boston: Houghton Mifflin Company, 1960.

Merton, Robert K. Social Theory and Social Structure. Glencoe, Illinois: The Free Press, 1957, revized and enlarged edition.

Mill, John Stuart. A System of Logic. New York: Harper and Brothers, 1847.

Nagel, Ernest. The Structure of Science. New York: Harcourt, Brace, and World, c.d., 1961.

Popper, Karl Raymond. The Logic of Scientific Discovery. London: Hutchinson, 1959.

_____. The Open Society and its Enemies. London: Routledge and Kegan Paul, Ltd., 1945.

Popper, Karl Raymond. The Poverty of Historicism. Boston: The Beacon Press, c.d., 1957.

Quine, Willard Van Orman. Word and Object. New York and London: John Wiley and Sons, Inc., c.d., 1960.

Ryle, Gilbert. The Concept of Mind. New York: Barnes and Noble, Inc., c.d., 1949.

Venn, John. Principles of Empirical or Inductive Logic. London and New York: Macmillan and Company, 1889.

Ziff, Paul. Semantic Analysis. Ithaca, New York: Cornell University Press, c.d., 1960.

Journals

Abel, Theodore. "The Operation Called Verstehen," American Journal of Sociology, 54:211-218, November, 1948.

Albery, Ethel M. "Causality in the Social Sciences," Journal of Philosophy, 51:695-706, November 11, 1954.

Bunge, Mario. "Do Computers Think?", Part I, The British Journal for the Philosophy of Science, 7:139-148, August, 1956.

_____. "Do Computers Think?", Part II, The British Journal for the Philosophy of Science, 7:212-219, November, 1956.

Feuer, Lewis S. "Causality in the Social Sciences," Journal of Philosophy, 51:681-695, November 11, 1954.

_____. "Indeterminacy and Economic Development," Philosophy of Science, 15:234, July, 1948.

Gewirth, Alan. "Can Men Change the Laws of Social Science," Philosophy of Science, 21:229-241, July, 1954.

Goodman, Nelson and Henry Leonard. "The Calculus of Individuals and its Uses," Journal of Philosophy, 5:45-55, June, 1940.

Grünbaum, Adolf. "Historical Determinism, Social Activism, and Prediction in the Social Sciences," The British Journal for the Philosophy of Science, 7:236-240, November, 1956.

Grünbaum, Adolf. "Temporally-Asymmetric Principles, Parity between Explanation and Prediction, Mechanism versus Teleology," Philosophy of Science, 29:146-174, April, 1962.

Hartshorne, Charles. "Freedom Requires Indeterminacy and Freedom," Journal of Philosophy, 55: 793-811, September 11, 1958.

Hempel, Carl G. and Paul Oppenheim. "Studies in the Logic of Scientific Explanation," Philosophy of Science, 15: 135-175, April, 1948.

Henle, Paul. "Do We Discover our Use of Words?" Journal of Philosophy, 54: 750-758, November 7, 1957.

Leonard, Henry S. "The Logic of Existence," Philosophical Studies, 7: 49-64, June, 1956.

Miles, T. R. "On the Difference between Men and Machines," The British Journal for the Philosophy of Science, 7: 277-292, February, 1957.

Quine, Willard Van Orman. "The Scope and Language of Science," The British Journal for the Philosophy of Science, 8: 1-17, May, 1957.

Rescher, Nicholas. "Belief-contravening Suppositions," Philosophical Review, 70: 176-196, April, 1961.

Richfield, Jerome and Irving M. Copi. "Deciding and Predicting," Philosophy of Science, 28: 47-52, January, 1961.

Roshwald, M. "The Case for Indeterminism," The British Journal for the Philosophy of Science, 7: 240-242, November, 1956.

Rudner, Richard S. "The Scientist Qua Scientist Makes Value Judgments," Philosophy of Science, 20: 1-6, January, 1953.

Scheffler, Israel. "Explanation, Prediction, and Abstraction," The British Journal for the Philosophy of Science, 7: 293-309, February, 1957.

Encyclopedias

Hempel, Carl G. "Fundamentals of Concept Formation in Empirical Science," International Encyclopedia of Unified Science, Vol. II, No. 7, Chicago: University of Chicago Press, c.d., 1952.