THE UNIVERSITY OF CALGARY

MODELS AND MODALITY

by

Stephen Downes

Submitted to the Faculty of Graduate Studies

In partial fulfillment of the requirements for the

Degree of Master of Arts

Department of Philosophy

Calgary, Alberta

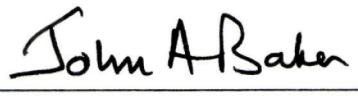
September, 1987

ISBN 0-315-37985-5

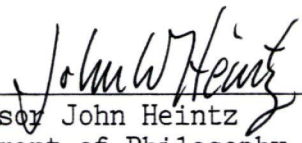
THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

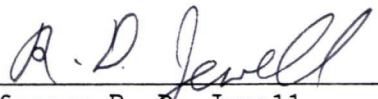
The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies for acceptance, a thesis entitled "Models and Modality", submitted by Stephen Downes in partial fulfillment of the requirements for the degree of Master of Arts.



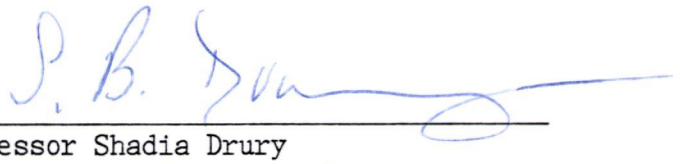
Professor John A. Baker, Supervisor
Department of Philosophy



Professor John Heintz
Department of Philosophy



Professor R. D. Jewell
Department of Philosophy



Professor Shadia Drury
Department of Political Science

September 10, 1987

ABSTRACT

Logical Positivism has been replaced by a new theory, a theory I call the model-modal theory. In this thesis I define the model-modal theory, show how it replaces logical positivism, and show why it fails.

In logical positivism the analysis and evaluation of statements was characterized by two theses: first, the analysis and evaluation was direct and based on the world in which the statement was asserted; and second, this analysis and evaluation occurred independently of the context of that assertion.

The model-modal theory rejects both theses. In their place it suggests that the analysis and evaluation of statements and propositions is indirect and accomplished with reference to some model of the world in which the statement or proposition is asserted. The selection of such a model from a set of possible models is accomplished with reference to the particular context in which the statement or proposition is asserted.

There are two major branches of the model-modal theory.

The first is what I call the 'possible worlds theory' which employs Kripke's possible worlds analysis of statements using the operators 'necessarily' and 'possibly' for statements using counterfactual operators such as 'could' and 'would'. Adherents

include Robert Stalnaker, David Lewis, Alvin Plantinga, Saul Kripke, and others.

The second is what I call the world view theory. On this theory scientific and other truths are determined within the context of some theory or world view. Philosophy of science should concentrate on the factors leading to the acceptance and rejection of such theories. Pioneered by Thomas Kuhn it has as followers N. R. Hanson, Michael Polanyi, Imre Lakatos, Larry Laudan, Bas van Fraassen, Philip Kitcher and many others.

In this thesis I show that both the possible worlds theory and the world view theory are instances of the model-modal theory by showing that they reject the two theses of positivism and employ a model structure as defined.

I identify the core idea and the core problem for the model-modal thesis as the selection of such models. To evaluate some statement or proposition which is asserted in some world G a model of that world H must be selected. But to select some model H it is necessary first to establish that some statements or propositions in G are true. This generates either a vicious infinite regress or a vicious circularity.

ACKNOWLEDGEMENTS

Professor John A. Baker, for insisting that I state everything precisely.

Professor John Heintz, for making modal logic so clear.

Professor Bob Jewell, for showing me why positivism fails.

Professor Gordon Greig, for showing me positivism still lives.

Damon Gitelman, for proofreading.

Karen Wyllie, for support and encouragement.

John Pugsley, Geoff Gorham and a host of others for listening in Dinnies.

The Coffee Company, for the many free coffees.

DEDICATION

To the land and people of the Province of Alberta.

Thank you.

CONTENTS

MODELS AND MODALITY

Chapter One

INTRODUCTION: THE MODEL-MODAL THEORY

1. PRELIMINARY REMARKS

A paradigm shift has occurred. It took place within the last thirty years, concurrent with the collapse of logical positivism. Positivism has been superseded. It has been replaced by what I shall call the model-modal theory.

Positivism was atomistic; statements were discrete entities, evaluated independently of other statements. The model-modal theory is holistic; no statement can be analysed without considering the totality of statements.

Positivism was unified; one and only one theory of knowledge or picture of the world would be the correct one for all disciplines. The model-modal theory is relativistic; a theory of knowledge or world picture can vary according to the context of the discipline in which it occurs.

Positivism employed a strictly deterministic logic; statements were either true or false, deductions either valid or invalid, without

variation. The model-modal theory employs a variably strict logic; statements can at one instance be true, at another, false, and deductions which were at one time valid may be at another time invalid.

Positivism was referential; statements which were not analytical were given meaning by reference to some object or objects specified by the discourse. The model-modal theory is representative; all statements are given meaning according to the manner in which the totality of statements describes or depicts some universe of discourse.

There are two major schools of thought, each apparently quite different from the other, which, when taken together, combine to compose the model-modal tradition. The first, pioneered by Kuhn, Hanson, and others, is what I will call the world-view theory. The second, pioneered by Kripke, Lewis, and Stalnaker, is what I will call the possible worlds theory. The first starts with model-theoretic structures. The second starts with modal theoretic structures. Each implies the other; they are necessarily conjoined.

So far as I am aware, no one has explicitly conjoined these two aspects of the model-modal theory in this way. Accordingly, a major part of this thesis is devoted to establishing that there is one unified paradigm which has emerged following the collapse of logical positivism. A second major part of my thesis is this. Each of the two aspects, when taken separately, appears to provide a strong and

tenable theory of knowledge and truth valuation. When taken together, however, it becomes possible to demonstrate that the model-modal theory is inherently viciously regressive or circular.

Or so I shall argue.

2. STRUCTURE OF THIS INTRODUCTION

This introduction has three objectives. First, in sections 3, 5 and 6 I sketch the development of the model-modal theory as it emerged from the collapse of positivism. Second, in sections 4, 7, 8, 9 and 10 I outline the core structure of the model-modal theory. And third, in section 11, I state for the first time the problems of vicious regress and circularity inherent in the model-modal theory.

I sketch the development of the model-modal theory as follows. First, in section 3 I identify a number of theories which emerged as a consequence of the collapse of positivism and divide them roughly into two groups: the 'world view theory' and the 'possible worlds theory'. Second, in section 5 I examine more closely the response of the world view theory to the failure of the observation / theory distinction. And third, in section 6 I examine more closely the modal semantical theory and its response to the failure of the analysis and valuation of counterfactuals.

I outline the core structure of the model-modal theory as follows. First, in section 4 I outline the core structure as a whole and show

briefly how it is employed by both the world view theory and the possible worlds theory. Second, in section 7 I define in some detail the model structure employed by the model-modal theory. Third, in sections 8 and 9 I argue that this model structure is in fact employed by both the world view theory and the possible worlds theory. And fourth, in section 10 I outline possible worlds semantics as it is employed in the model-modal theory.

I state the problem of regress and circularity in the model-modal theory for the first time in section 11. In particular I show how these problems occur when the model-modal theory is employed to solve the problems identified as the cause of the collapse of positivism. In section 12 I show how the argument developed in section 11 will be developed through the course of this thesis.

3. THE DEVELOPMENT OF THE MODEL-MODAL THEORY: AN OUTLINE

Positivism was rejected for three major reasons: first, the observation-theory distinction proved to be untenable; second, no satisfactory analysis of counterfactuals was possible; and third, methods of confirmation failed to provide a sufficient validation or vindication of the assigned truth values of propositions.

Several apparently distinct theories emerged as a result of this collapse. They may be briefly sketched as follows.

First, Kuhn, Hanson and others, in response to the failure of the observation-theory distinction, argued that observation statements are to be understood only in the context of a paradigm or world view. I shall call this the world view theory.

Second, Lakatos, Laudan, and others argued that there is to be no evaluation of various world views or theories based on the merits of the theories themselves; philosophy should therefore concentrate on the social and external factors which motivate the acceptance and rejection of theories and world views. This (it seems to me) is a logical outgrowth of the world view theory.

Third, Quine, Goodman, Chisholm and others, after considering the failure of the analysis of counterfactual and causal statements, argued that such statements are to be understood and analysed only in terms of possible worlds or world views. In this view are elements of the world view theory and the possible worlds theory.

Fourth, Kripke and (a very few) others developed a possible world semantics. This is an essential component of the possible worlds theory.

Fifth, Stalnaker, David Lewis and others applied Kripke's possible worlds semantics to the analysis of counterfactual and causal propositions such that the truth values of statements expressing counterfactual or causal propositions are to be determined with reference to possible worlds. This is the possible worlds theory of counterfactuals (sometimes called the Lewis-Stalnaker theory).

Sixth, a constructivist program based largely on the work of Kuhn and Hanson was developed by Kitcher and van Fraassen. This development is largely a refinement of the world view theory and for the most part I shall treat it as such.

In my discussion below I shall divide these varied approaches roughly in two. The first is the world view theory and shall be described in some detail in this introduction and then in some greater detail in chapters four and five. The second is the possible worlds theory and shall be described in outline form in this introduction, some greater detail in chapter two, and with particular attention to its analysis and valuation of propositions in chapter three.

Each of these theories employs models. And, perhaps more important, each employs a multiplicity of models. To establish this it is necessary to distinguish between the essential and accidental features of models. I shall argue that a model may have any ontological status and give examples of three distinct such ontological statuses. Secondly, I shall argue that a model has as an essential quality a relation to something which is not a part of the model itself. I shall, using these considerations, prove that both possible worlds and world views are models.

4. THE MODEL-MODAL THEORY: AN OUTLINE

My discussion of the core structure divides roughly into three parts. In the first part I define in some detail a model structure

based roughly on a Kripke model structure and our ordinary use of the word 'model'. This occurs in section 7. In sections 8 and 9 I demonstrate that the model structure described in section 7 is employed by both the world view theory and the possible worlds theory. And third, I outline possible world semantics as employed by the model-modal theory in terms of the model structure defined in section 7.

I suggest two crucial theses in section 7. First, I suggest that for an entity to be a model is not to be an entity of a certain type or ontological status. Second, I suggest that for an entity to be a model is for that entity to be used as a model. To be used as a model, I suggest further, is to be shown to be related in a particular way to some other entity which I call an 'original'.

In section 8 I argue that a 'possible world' in the possible worlds theory is in fact a model according to the definition offered in section 7. In section 9 I argue that a 'world view' in the world view theory is also a model according to that same definition.

In section 10 I outline the possible world semantics employed in the model-modal theory. I suggest that some relation R is employed to successively restrict the number of possible models which may be used as a model of some original. I also show how this successive restriction is integral to the procedure of determining the truth value of some proposition of the original with reference to the truth value of the same proposition in some model of the original.

Although I will go to great lengths below to show that both the world view theory and the possible worlds theory employ this common core I would like at this point to suggest some reasons why we at this point might consider the idea initially plausible.

First, both theories employ a model which is intended to represent the world as if it were a certain way. It is true that the ontological status of this model varies from theory to theory but a model as I have defined it may be of any ontological status. Rather, what is important is the use of some entity as a model. By use I mean that some relation is shown to hold between the model and the original. Possible worlds are related by similarity to the actual world in the modal semantical theory. And world views (or theories) are related by descriptive similarity to the world they are intended to describe.

Second, in the model structure outline in section 7 there is a multiplicity of possible models for any original and some process of selection is required, as detailed in section 10, to select for use some particular model. In both the possible worlds theory and the world view theory the selection of a model (possible world or world view) occurs.

Third, in many versions of both the modal semantical theory and the world view theory the criteria for this selection are the same. The model selected must be possible, that is, non-contradictory. And it must be related to the original in such a way that salient facts

(salience determined by the context of use) are true in both the model and the original.

Fourth, in many instances a proposition is true in the original world if and only if it is true in the model or world view. An exact parallel exists in some theories (such as van Fraassen and Hanson) about observation statements. A statement is an observation statement if and only if in the selected model (world view) it is an observation statement.

5. THE COLLAPSE OF THE OBSERVATION-THEORY DISTINCTION AND THE RISE OF THE WORLD VIEW THEORY

In logical positivism, as most of us are aware, general laws or axioms were postulated or perhaps inferred from observational data. From these laws observation statements or propositions were deduced. The general laws or propositions were in turn confirmed or disconfirmed according to whether the deduced propositions correctly described subsequent observations. The failure of positivism, of course, occurred largely as a consequence of philosophers' inability to specify exactly how this procedure would take place. No effective distinction could be drawn between observation statements and other components of the theory in question.[1] In the absence of such a

---

[1] Quine's 'Two Dogmas of Empiricism' initiated the process. Putnam's 'The Analytic and the Synthetic' and Achinstein's 'The Problem of Theoretical Terms' delivered what Suppe called the "coup de grace".

reduction it was impossible to conclusively verify or falsify any scientific theory.

I am not so concerned to discuss the reasons for the collapse of the observation - theory distinction; I think they are well known. More important is the consequence. There has occurred a recognition by philosophers on all sides that what is to count as 'observable' or as an 'observation statement' will have to be defined by the theory in question itself. Statements which are considered within the context of the theory to be observation statements are now considered to be irreducibly theory-laden. This approach, which was pioneered by Hanson, Polanyi, Kuhn and others, has been well documented and I will not dwell on it at length. What is key to the new approach is the thesis that what one observes is constituted in part by the Weltanschauung the "disciplinary matrix" or the "world view" of the observer.[2] That is, what one observes and what one theorizes is all part of a complex whole.

I do not want to suggest that a single theory has been proposed. Approaches which follow the Weltanschauung analysis vary from the radical context dependence proposed by Hanson to the realism of Feyerabend and Popper. It is rather the general thesis that our understanding of science and knowledge must occur from within a previously defined conceptual framework which I wish to emphasize. I think Suppe best summarizes the new approach:

---

[2] These terms are taken from Suppe, "The Structure of Scientific Theories", pp. 125-190.

> What is required is an analysis of theories which concerns itself with the epistemic factors governing the discovery, development, and acceptance or rejection of theories; such an analysis must give serious attention to the idea that science is done from within a conceptual perspective which determines in large part which questions are worth investigating and what sorts of answers are acceptable; the perspective provides a way of thinking about a class of phenomena which defines the class of legitimate problems and delimits the standards for their acceptable solution.[3]

6. THE FAILURE TO ANALYZE COUNTERFACTUALS AND THE RISE OF THE POSSIBLE WORLDS THEORY

Concurrent with the collapse of the observation - theory distinction came the collapse of standard analyses of statements intended to describe or document inductive or causal hypotheses. Although hints of this problem date back to Hume's Treatise the standard statement in terms of what we now call counterfactual conditionals was first provided by Goodman and Chisholm in the late 1940's.[4] Laws of logic which apply to the standard conditional seem inappropriate or just wrong for the counterfactual conditional. To understand the proposed solution to the problem it is necessary to pinpoint more accurately exactly where the problems occur. I shall do that briefly before moving on.

First, the normal laws of contraposition and transitivity, which hold under the standard analysis, do not seem to be appropriate for

---

[3] Suppe, p. 126.

[4] In particular, in Chisholm's 'The Contrary to Fact Conditional' and Goodman's 'The Problem of Counterfactual Conditionals'.

counterfactuals. For example, if we assert 'if the match had been struck it would not have lighted' we do not normally intend also to assert the contrapositive 'if the match had lighted, it would not have been struck'. As a second example, if we assert both 'if Hoover was a communist he would have been a traitor' and 'if Hoover was a Russian he would have been a communist' we do not necessarily want to assert 'if Hoover was a Russian he would be a traitor'.[5]

Second, the standard conditional analysis does not seem sufficient to distinguish between causal generalizations, accidental generalizations, and a myriad of possibilities in between. For example, if we heat water, it boils; that is, little bubbles form and steam rises. We say that the heat causes the steam and the bubbles on the grounds that whenever water is heated sufficiently these phenomena occur. However, on the same grounds, we might say that the rising of the steam causes little bubbles to form and heat to occur.

There is an important distinction between the first and second problem. The first is a problem of the analysis of counterfactuals. It is a problem of specifying the procedure to be employed in the determination of truth values of a counterfactual. The second is a problem of the evaluation of a counterfactual. It is a problem of determining whether the content of the counterfactual is 'correct' or 'true of the world'. The first problem is addressed with a possible

---

[5] The first example is from Goodman, op. cit. and "Fact Fiction and Forecast", p. 5. The second is from Stalnaker, 'A Theory of Conditionals' and reprinted in "Ifs", p. 48.

worlds analysis of counterfactuals. The second problem is addressed with a possible worlds evaluation of counterfactuals.

On the model-modal theory the solution to both problems rests on the suggestion that counterfactual propositions should be understood as obtaining truth values from within the context of a possible world. Exactly what a 'possible world' is varies from theory to theory. I wall argue below that possible worlds are best viewed as models. But for now let me examine how possible worlds are employed to solve both the analytic and evaluative problems of counterfactuals.

First, the analytical problem is resolved by the recognition that the laws of contraposition and transitivity are not in all cases applicable to counterfactuals. Possible worlds are employed to explain why these laws are not applicable and to provide a procedure for the determination of counterfactual truth values. Counterfactuals are true or false relative to possible worlds. Counterfactual truth is variable because different possible worlds may be selected in different circumstances.

Second, the evaluative problem is solved with the employment of possible worlds semantics. Possible world semantics employs the analysis of counterfactuals in terms of possible worlds as explained immediately above. In addition it employs what is called a 'selection function' which will be used to choose the possible world with which the truth value of a particular counterfactual will be determined.

The selection of a possible world will be determined by features of the counterfactual statement and the context in which it is asserted.

Neither an analysis of nor the evaluation of counterfactual conditionals may be determined, according to the model-modal theory, without the use of possible worlds. The fact that there are possible worlds allows for a variably strict analysis of counterfactuals. And the content of such possible worlds will determine the content of true counterfactual statements in the actual world.

I shall examine the analysis and evaluation of counterfactuals in much greater detail below. First, however, it is important to understand the use and structure of models in the model-modal theory. I shall subsequently explain the modal semantics of counterfactuals and then examine in greater detail how this semantics is employed in the determination of truth values for different propositions.

7. MODELS

A concept common to both the world view theory and the possible worlds theory is a concept I shall express using the word 'model'. In this section I shall explain exactly what I mean by a 'model'. In the next two sections I shall argue that both the world view theory and the possible worlds theory use this concept in an important way.

My use of the word 'model' to express this concept, while stipulative, is not entirely arbitrary. My usage will have close

affinities with both the ordinary language use and Kripke's formal use of the word. But my use of the word is a new use and exact correspondence with either Kripke's use or ordinary use is neither expected nor desired.

To be a model is not to be an entity of a particular type or ontological status. Rather, to be a model is to be an entity with is put to a particular purpose. For example, a pile of pebbles is just a pile of pebbles and not a model of the solar system until the pile of pebbles is used as a model of the solar system.

To be used as a model an entity must be shown to be related in some way to some distinct entity which I shall call the 'original'. An original, like a model, need be of no special type or ontological status. To be an 'original' is to be an entity to which a model is intended to be related in some way.

We can use Kripke's 'model structure' to represent more formally the concept introduced above. "A model structure (m.s.) is an ordered triple (G,K,R) where K is a set, R is a reflexive relation on K, and G (is a member of the set) K." Using the terms introduced above, G is the 'original', K is a set of possible models, and R is the relation between the model and the original.

On Kripke's usage, G is intended to be the 'real world' and the relation R is intended to define the extent of the set K as 'possible relative to G'. On my usage G need not be the 'real world'. An original need be of no particular ontological status, thus G may be

'non-real'. It is best to view G as a member of the set K, as above. That is, it is best to view G as a model itself.

As I have suggested, an entity is a model if it is shown to be in some relation to an original. It is crucial to specify more precisely the nature of this relation. We may do so by examining more closely the nature of the relation R on the modified Kripke model structure.

The relation R may serve three functions. First, R may limit the size of K by restricting K to entities which are possible relative to G. Second, R may select some G out of the set K such that G acts as an original. Third, R may select some member of K which I shall call H as a model such that H is a model of an original G.

K may be considered as the set of all possible models. Since all models are possible relative to themselves there is no need to restrict this to possibility relative to some original G. That leaves R with two purposes: the selection of some G of K, and the relation of this G to some H of K. But we cannot eliminate one of these purposes without eliminating the other. It will be impossible to select some G from K without also relating it to some H. It will be impossible to relate some H to G without first selecting some G.

Let us therefore define a model H as follows. A model H is an entity such that (i) H is possible relative to itself, (ii) H stands in relation R to some G which is also possible relative to itself such that R selects G as an original relative to H, and (iii) H stands in relation R to some G such that R selects H as a model of G.

The relation R has not been thus far defined as a certain type of relation; rather, it has been defined only as a relation which accomplishes a certain purpose (to wit, the selection of an original G and a model H). I do not propose to restrict R to a certain type. Nonetheless, it will be useful to describe R as it is most commonly defined: as a similarity relation.

We typically use some entity H as a model of some other entity G if H is in some way relevantly similiar to H. But H may be similiar to G in a variety of ways. I shall indicate briefly four distinct ways in which H may be similiar to G.

The first pair of ways defines the types of similarities between some G and some H. First, the same atomic formulae might be true of both G and H. For example, both G and H might be red, or round, or have seeds, and so on. G and H might be qualitatively similiar. Second, the same n-tuple relations might be true of both G and H. For example, both G and H might hang from trees, bounce when dropped, and so on. G and H might be relationally similiar.

The second pair of ways defines the mode of similarity between G and H. First, H might be directly similiar to or isomorphic with G. For example, a picture of a mountain will be similiar in this way. A one-to-one mapping of some parts of H to some parts of G is possible. Second, H might be indirectly similiar to or descriptive of G. For

example, a description of a mountain will be similiar to a mountain in this way. No one-to-one mapping is possible.[6]

Whether a similarity relation may be employed as a relation R which defines some entity H as a model of some other entity G does not depend on the type or mode of similarity employed. Any type or mode will do at one time or another. What is crucial is that the relation R be such that some H may be designated as a model and that some G may be designated as an original.

8. MODELS AND POSSIBLE WORLDS

In this section I want to argue that possible worlds are models. By that I mean that the set of all possible worlds is a proper subset of all possible models. The set of all possible worlds itself divides into several subsets. One such subset is the set of all real possible worlds. Another is the subset of all ersatz possible worlds. And so on. Any possible world, regardless of its ontological status, is a model.

It might seem odd at first to consider a possible world to be a model. David Lewis, for example, considers the possibility and then rejects it. But I suggest that Lewis rejects it because he has used too narrow a definition of 'model'.

---

[6] One might say that the similarity occurs between the meaning of the words employed in the description but that, I think, would be to insist on a definition of similarity which must be isomorphic. No such restriction need be employed.

Lewis suggests that possible worlds might be "maximal consistent sets of sentences in some language... that is state descriptions; or... diagrammed models."[7] He then rejects the suggestion on the grounds that "I emphatically do not identify possible worlds in any way with respectable linguistic entities."[8]

The fact that, on Lewis's account, possible worlds are not linguistic entities does not, on my account, preclude them from being models. Some models are linguistic entities. But not all models are linguistic entities. Some models are as real as Lewis's possible worlds. No entity can be precluded from being a model because of its ontological status because, as I have argued above, a model may be of any ontological status.

We may assert that possible worlds are models because they are defined in the manner models were defined in the previous section. The set of all possible worlds is the set K in the model structure (G,K,R). The real or actual world is G and G is a member of K. A possible world is some member of K which is possible relative to itself (otherwise it would be an impossible world). G is actual with respect to H. H stands in relation R to G according to a similarity ordering.

The precise manner in which some G is selected as the real or actual world will be the subject of detailed discussion below. The

---

[7] Lewis, "Counterfactuals", p. 85.

[8] Ibid.

precise manner in which some H is defined as similiar to G will also be closely examined. I shall therefore let the discussion below serve as proof for the preceding paragraph.

9. MODELS AND WORLD VIEWS

It is more difficult to describe world views as models because there is no obvious sense in which some particular world view is selected as an 'original'. The central thesis of the world view theory is that there is no 'real' world which is described by a theory but rather only a number of possible theories. That is, the world view theory, properly understood, would use not the triple (G,K,R) but rather the singlet (K).

We should at this point distinguish between two major branches of the world view theory. First: there is an real world but it may be described by any of a number of theories none of which is uniquely 'correct'. Second, there is no original world but only a number of theories none of which is uniquely 'correct'. Either way, the selection of which theory to employ in any given circumstance will not depend on the 'world' per se but rather on other factors.[9]

On the first case world views may be fairly easily defined as models. The real world is some G of K and the various theories are other members of K which are models of G: call them H1, H2, and so

---

[9] For example, Laudan cites 'problem solving', Lakatos suggests socially determined 'research programs', and so on.

on. Nothing in my definition of models above suggested that G must be completely described or uniquely determined; the theory requires only that there be some G. H1, H2, and so on are related to G in the sense that they are intended to describe G. That neither H1 nor H2 serves as a unique description of G is of no account. G in any case cannot be specified without recourse to some H1 or H2 and we might suggest that for each H1, H2, and so on there is a unique G1, G2, and so on. Since each theory is a world view, each world view is a model.

A similiar argument may now be advanced with respect to the second version with one proviso: nothing in my theory suggests that G must be real. The 'original' G might be some world as described by some theory T. Thus, for each theory T there will be a corresponding G such that G1 is described by T1, G2 by T2, and so on. Each T is therefore a model of some G; and since each theory is a world view on the second version as well, each world view is also a model.

It would be a mistake to think that on the model theory I am describing that the original G must be something unique, real and distinct from the models in the set K. In fact, on this theory, any member of K might be, from the standpoint of some H, G. To be sure, many theorists want to select as G some member of K which is both real and actual; but they all have in common some means of selecting a particular G from K and beyond that very little in common at all.

It would also be a mistake to think that some entity H which is a description of some entity G cannot be a model of G. That, I suggest,

would be to arbitrarily restrict the relation R to a direct, isomorphic mode. Philosophers working within the model-modal theory might consider the isomorphic versus description question to be of extreme importance. All I need assert is that both sides of that debate nonetheless fall within the tradition.

10. POSSIBLE WORLD SEMANTICS

Both the possible worlds theory and the world view theory employ possible world semantics. By possible world semantics I mean most generally the following: the use of some model H to determine the truth value of some proposition O(A->B) of an original world G.

For the most part modal semantics has been most completely developed within the possible worlds theory. A large part of chapter 4 is devoted to demonstrating that it is also employed in the world view theory. In this section I shall describe possible worlds semantics as developed in the possible worlds theory and only briefly show how it may be proven that it is also employed in the world view theory.

I shall use the following notation to designate a proposition: O(A->B). The use of the letter 'O' is intended to indicate that the expression O(A->B) denotes a class of statements of the type A->B. For example, a statement of the type A->B might be, "this match, when struck, lights". The corresponding proposition would then be, "matches, when struck, light".

Let me state now in barest form the possible worlds theory: a proposition O(A->B) is true in G if and only if (i) there is some H such that H is most similiar to G, and (ii) O(A->B) is true in H. On the possible worlds theory, G refers to the actual world and H refers to some possible world.

The possible worlds theory may be placed within the model structure outlined above: (G,K,R). G is the actual world and a member of the set K, the set of all possible worlds. The relation R is employed to select some H of K which is relevantly similiar to G.

There are three distinct stages to the selection of some H on the possible worlds theory. They are as follows:

First, R is employed as an 'accessibility relation' which defines which members of K are possible relative to G. In the possible worlds theory as stated by Lewis and Stalnaker this is the only explicit use of R; the remaining two steps are defined separately.

Second, of those members of K allowed by the first step a selection function f chooses the world or worlds in which the antecedent of the proposition A is true and which is the closest (most similiar) to G given the truth of A.

Third, if more than one world has been selected by the second step (or if the selection is relevantly unsatisfactory) then a more rigid designation of the qualities of G to which H must be similiar is

defined by the context in which the truth valuation occurs. These qualities of G are called 'salient qualities'.

Although, as I stated, R is used only in the first of these three steps by Lewis and Stalnaker I think that we should more properly consider R to encompass all three steps. Each step is the successive restriction of those members of K which will be used as a model of G in the process of determining the truth value of some proposition in G. The steps most properly ought to be considered as the successive refining of R as possibility, similarity, and relevant similarity.

The evaluation of propositions in the world view theory occurs in a parallel manner. A proposition P is true in G described by T if and only if it is true in T. In cases where there are competing T's for one G (as in the first type of world view theory described above) the proposition is true in G if and only if it is true in some T which is selected in much the same manner a possible world is selected in the three stages outlined above. In cases where there is one G for each T then only one T is possible relative to G (presumably) and hence only the first stage above is needed.

In chapters four and five I show how both the theories of Kitcher and van Fraassen evaluate propositions in this manner. But it is more widely applicable in the world view theory. For example, for N.R. Hanson, what is 'observed' will depend on the world view held by the observer. That is, the value of a statement as an 'observation statement' depends on the theory in which it is asserted. The theory

is a model of the world in which the observation is supposed to have occurred. To say, in Hanson's theory, "I saw the sun rise" is to say "I saw such and such an 'observation' in such and such a world as described by such and such a theory."

11. PARADIGM LOST: PROBLEMS FOR THE MODEL-MODAL THEORY

The model-modal theory was intended to solve the problems which caused the collapse of logical positivism. In this section I want to outline some problems which, if not resolved, will cause the model-modal theory itself to collapse.

In broad outline, the problems which caused the collapse of positivism remain unsolved by the model-modal theory. This may be demonstrated by a two stage argument. In the first stage I suggest that the analysis of counterfactuals offered by the model-modal theory cannot be employed as intended without vicious circularity. In the second stage I argue that the semantics employed by the theory cannot determine the value of any statement without either vicious regress or circularity.

The analysis offered by the model-modal theory is intended the replace the failed analysis of counterfactuals proposed by logical positivism. The semantics offered by the model-modal theory are intended to determine, as positivism could not, when a particular proposition is a true proposition and when a particular statement is an observation statement, that is, to properly evaluate propositions.

In this section we take as given the model-modal theory as described above. In particular, we take as given the definition of a model and the definition of similarity outlined in section 8 and the definition of modal semantics as outlined in section 10.

Let me now present the first stage of my argument. I suggest the following. The model structure employed by the model-modal theory cannot be used to analyze counterfactuals without infinite regress. Recall that the model structure is an ordered triple (G,K,S). The first stage of my argument concludes as follows: it is impossible to specify from the set K any G or any H which is a model of G without infinite regress. The argument is a four part argument. I shall outline each part.

First. H is the closest possible world to some G. Which member of K is H cannot be determined until we specify some G (or part of G) which H would be closest to. Therefore we need G to find H.

Second. G needs to be specified. But G cannot be specified without first specifying some model of G (which is H) because any specification of G will include at least in part the assertion that some counterfactual proposition is true in G. Therefore we need H to find G.

Third. If we attempt to specify some G then we must specify some H. But any attempt to specify H will in turn require that we specify some G. Therefore we cannot begin by attempting to specify some G.

Fourth. If we attempt to specify some H then we must specify some G. But to specify some G we must specify in turn some H. Therefore we cannot begin by specifying some H.

That concludes the first stage of the argument. The modal analysis must fail because we cannot specify any original G or model H of G without vicious circularity. This form of the argument is suggested in chapter two.

The second stage of the argument applies the first stage to specific instances of evaluation. I shall now turn to the second stage of the argument.

By 'evaluation' I mean the specification of some particular value for a particular proposition. In the second stage I argue that it is impossible to specify the value of any proposition which requires modal semantics for this determination.

There are three types of value a proposition may have. First, a proposition may or may not be true. Call this value a truth value (t-val). Second, a proposition may or may not be asserted with an observation statement. Call this value the observation value (o-val). Third, a proposition may or may not be a causal proposition. Call this value the cause value (c-val).

In my statement of the second stage of this argument I state it in terms of truth values for counterfactual propositions (c-p) throughout. By substituting all occurrences of t-val with o-val and

by replacing all occurrences of c-p with 'proposition' an exactly parallel argument may be constructed for observation values. A similiar parallel may be constructed for c-vals. This second stage has nine steps.

First. Suppose we want to find the t-val in G of some c-p. To do this we must find some H which is closest to G.

Second. Closeness is defined by R. To find out how close some H is to some G we must find the truth values of some set (defined by R) of atomic formulae (a-f) or n-tuple relations (n-t) in both G and R.

Third. In some cases the a-f or n-t which must be given t-vals in G and H will be c-p.

Fourth. We need in some cases to find the t-val of some c-p in either G or H.

Fifth. To find the t-val of some c-p in G go to the first step. (Vicious circle; no t-val in the original case can be determined.)

Sixth. To find the t-val of some c-p in H we need to find some member L of the set K (from (G,K,R)) which is closest to H.

Seventh. Closeness is defined by R. To find out how close some L is to some H we must find the truth values of some set (defined by R) of a-f or n-t in both L and H.

Eighth. To find the truth values of some c-p in H go to the sixth step. (Vicious circularity, no t-val established).

Ninth. L is either G or not G. If L is G then go to the first step (vicious circle). If L is not G then go to the sixth step substituting L for H. (infinite regress).

No matter which procedure is followed anyone who attempts to find the t-val of some c-p in either G or H will encounter either an infinite regress, or vicious circle.

This second stage of the argument is also stated in detail in the chapters below. In chapter three the first to sixth steps are established. In chapter four the sixth through ninth steps are established.

This concludes the second stage of the argument. Before continuing it is best to briefly state exactly what I mean by 'specify the truth value of' or 'specify G'. By 'specify' I mean to determine in any way whatsoever. In particular I mean to cover by 'specify' two distinct means of determining what some truth value is or which world of K is G.

First: stipulation. If some G or some t-val in G is determined by prior stipulation then a certain H must also be determined by prior stipulation. Circularity or regress in this instance means that, once started, such stipulation may never cease and hence never succeed in determining some G or some t-val in G.

Second: inference. If some G or some t-val in G is determined by some process of inference then a certain H must also be determined by

that process of inference. Circularity or regress in this instance means that, once started, such inference may never cease and hence never succeed in determining some G or some t-val in G.

12. STRUCTURE OF THIS THESIS

This concludes the introductory chapter of this thesis. The next three chapters of this thesis will be devoted to proving that problems of the sort described in section 11 do in fact exist in the model-modal theory

In this introduction I have stated the following. Logical positivism has been replaced by a new paradigm I called the model-modal theory. This new paradigm emerged in response to the problems which caused the failure of logical positivism. Although the model-modal theory is apparently two distinct theories at the core of both are a model structure as described in section 7 and modal semantics as described in section 10. But in section 11 I argued that the model-modal theory faces severe problems of regress and circularity.

In chapter two of this thesis I shall closely examine the modal analysis of counterfactuals. It will be my intention in this chapter to show that this analysis employs a model structure as described in section 7 and modal semantics as described in section 10. In addition I will point the manner in which this theory was employed to solve the

problems which plagued logical positivism. The chapter concludes by suggesting the first stage of the problem outlined in section 11.

In chapter three of this thesis I shall substantiate the portions of the argument indicated in section 11. I shall establish the first to sixth parts of the second stage. In addition I shall continue to show how the argument is intended to solve the problems which plagued logical positivism and use as an example the particular problem of the analysis and valuation of causal propositions.

In chapter four of this thesis I shall substantiate the portions of the argument indicated in section 11. I shall establish the sixth to ninth parts of the second stage. In addition I shall show how the model-modal theory is intended to solve the problem of the observation-theory distinction and point to exact parallels between this solution and the modal analysis of counterfactuals outlined in the previous two chapters.

In chapter five of this thesis I examine the world view theory's attempts to escape the problems of circularity and regress describe in section 11 by recourse to theories of pragmatism and pragmatics. I detail such theories and show why the problem still occurs in all such instances of the world view theory.

Chapter Two

THE POSSIBLE WORLDS ANALYSIS OF COUNTERFACTUALS

1. INTRODUCTION

"If kangaroos had no tails they would topple over." We understand this counterfactual sentence. It has a meaning. It is possibly true or false. And even if all kangaroos do in fact have tails, we understand this sentence as saying something about the real world, this world, the one in which you and I are conducting this discussion.

But to explicate this type of sentence, to explain what we mean, and what we don't mean, by sentences using the conditional operators 'would' and 'could', is not straightforward. How, for example, do we verify such a sentence given that there are no kangaroos without tails. How do we construct a general logical form for such a sentence which does not violate our general intuitions about what is implied, and what is not implied, when we make such an assertion?

Logical positivism collapsed in the face of such questions and a primary factor in the rise of the model-modal theory was the desire to answer such questions. In this chapter I examine why positivist

responses failed and how the model-modal theory responds to this failure.

I have two theses to establish. First, the possible worlds analysis of counterfactuals employs a model structure as defined in chapter one. And second, the possible worlds analysis of counterfactuals has as its core thesis and core problem the selection of one or another model H of an original G from the set K.

This chapter has six sections of which this introduction is the first. In section 2 I outline the collapse of the positivist analysis of counterfactuals. In particular I identify two prevalent theses which the positivists could not sustain: (i) that counterfactuals may be analysed by direct inference from the world in which they are asserted, and (ii) that the analysis of counterfactuals is not dependent on the context in which they are asserted.

In section 3 I outline the possible worlds response to this failure in terms of two key and contrasting theses: (i) that counterfactuals are not analysed in isolation but rather with respect to the context in which they are asserted, and (ii) that true counterfactuals are not directly inferred from the world in which they are asserted but rather by indirect inference from some possible world or model which is relevantly similiar to the world in which the counterfactual is asserted. In this section I suggest that the core thesis and problem for the possible worlds analysis is the selection of such a model or possible world.

In section 4 I examine Stalnaker's definition of a 'selection function'. I show how Stalnaker's analysis fits into the model structure described in chapter one and describe the 'rules' or 'conditions' employed to select some possible world H from the set K on the basis of a similarity with a base world G. I show that Stalnaker's selection function must assume that the base world G must be described in some way and suggest that such a description may require that some counterfactual truths are already known.

In section 5 I examine the accessibility relation R which restricts membership of possible worlds on the basis of their 'possibility relative to G'. I show how the accessibility relation is employed by Kripke as R in the model structure (G,K,R) and explain how it is used in counterfactual semantics. The accessibility relation, I suggest, can be used to replace Stalnaker's selection function and thus apparently solves the problem inherent in the selection of some H similiar to some G.

In section 6 I show how the accessibility relation R is defined by restricting membership in K to worlds in which some salient qualities are present. In turn, I show how such salient qualities are defined by the context in which a counterfactual is asserted. I show how salience can order possible worlds in terms of their similarity with G and thus functions in the same manner as the selection function and the accessibility relation. But the same problem occurs: the determination of salient qualities requires a description of at least

some part of G which may involve the assertion that some counterfactual statement is true.

2. ANALYSES OF COUNTERFACTUALS IN THE POSITIVIST TRADITION

In the positivist tradition analyses of counterfactuals are characterized by two major theses.

First. Counterfactuals may be correctly analysed in isolation; that is, independently of the context in which they are asserted.

Second. Counterfactuals may be correctly analysed as variations of material or strict conditionals such that a true counterfactual statement may be inferred directly from evidence present in the world in which it is asserted.

In this section I shall demonstrate the collapse of both theses. In the next section I shall describe the development of two contrary theses which are central to the model-modal theory, specifically:

First. Counterfactuals are not analysed in isolation but rather are analysed from within the context of a possible world.

Second. The possible world to be employed is not described by direct inference from the world in which the counterfactual is asserted but rather is selected on the basis of similarity with the world in which the counterfactual is asserted.

For the remainder of this section I shall outline the collapse of the two positivist theses. There is a large volume of writing on this subject and therefore this account will be necessarily sketchy. Further, I am concerned for the moment only with the analysis of counterfactuals. I shall consider the evaluation of counterfactuals in the next chapter.

I shall sketch the collapse of these two theses as follows. First I shall explain why counterfactuals are not obviously material or strict conditionals. Second, I shall outline the positivists' use of an 'extended antecedent' in an attempt to save the analysis. Third, I shall explain why the 'extended antecedent' solution did not work.

Unless analysed as material or strict conditionals true counterfactual conditionals cannot be inferred directly from evidence in the world in which they are asserted. But it soon became clear that such an analysis would involve some complications.

At first glance the counterfactual conditional is very different from the material conditional. Material conditionals are always true if the antecedent of the conditional is false. Counterfactual antecedents are always false. But not all counterfactual conditionals are true.

Strict conditionals assert the necessity of the corresponding material conditional. The antecedent is always assumed to be true; the strict conditional is true only if the consequent is necessarily

true given that the antecedent is true. But this does not always apply to counterfactuals.

First, for some true counterfactuals the consequent is not necessarily true when the antecedent is true. For example, consider the true counterfactual "If the water had frozen the pipes would have broken." If it is true that the water had frozen it is not necessarily true that the pipes would break. There might be very little water in the pipes or the pipes might be very strong. The counterfactual should therefore be false but there are at least some instances where we would want it to be true.

Second, in some cases where the consequent is necessarily true given that the antecedent is true the corresponding counterfactual may be false. The intuition is that there should be some connection between the antecedent and the consequent and yet the consequent may be necessarily true no matter what the consequent may be.

For example, suppose the counterfactual, "If China had entered the Vietnam war the United States would have used nuclear weapons." Suppose further that the United States did in fact use nuclear weapons (or would inevitably have used nuclear weapons) even though the Chinese would never have entered Vietnam for any reason. Our intuition is that the counterfactual is false even though on the strict conditional analysis it must be true.[10]

---

[10] This example is borrowed and updated from Stalnaker, 'A Theory of Conditionals', p. 100.

The standard positivist response is to suggest that the material or strict conditional analysis would be adequate if implicit premises conjoined with the antecedent were explicitly stated. More formally, suppose the counterfactual

A -> B

is false but determined to be true by a material or strict implication. Explicitly stating the hidden assumptions as follows:

A & (C1 & C2 & ...) -> B

would solve the problem.[11]

For example, suppose I assert counterfactually that "If I had struck this match it would have lighted". If I assume for the sake of argument that the antecedent A is true I must also assume a number of 'cotenable' hidden premises: "the match was well made", "I struck it correctly", and "the match was not wet", to name a few.[12]

Analyses in the positivist tradition thus turned to an examination of exactly how to determine the set of cotenable premises. Many methods were proposed but all had in common the same general idea.

---

[11] The extended antecedent is called the 'set of relevant conditions' by Goodman in 'The Problem of Counterfactual Conditionals'. The same idea is employed slightly differently by Chisholm and Mackie, to name a few. The elliptical presentation of (C1 & C2 & ...) was, I believe, first introduced by Mackie (although he presented it slightly differently as x1,...,xn). This proposal is called the "Metalanguage" proposal by Lewis since Goodman, et. al., felt that linguistic convention could be employed to determine the content of extended antecedents.

[12] This example is from Goodman, 'The Problem of Counterfactual Conditionals'.

First, assume some set or 'stock of beliefs' C1 & C2 & ... about the world. Second, assume that some counterfactual antecedent A is true. Third, adjust the 'stock of beliefs' as necessary given the assumption of A. Fourth, determine whether the consequent B is a member of that revised stock of beliefs. If B is a member, the counterfactual is true; otherwise, it is false.[13]

Exactly how the stock of beliefs was to be 'adjusted' varied but all positivist theories had at least one of the following two features in common. First, a 'new' belief was added as a deductive or highly probable consequence of the assumption of A. Second. An 'old' belief was deleted as contradictory to or highly improbable given A. In both cases, the adjustment of the stock of beliefs could be determined strictly based on the assumption of A.[14]

The positivist 'extended antecedent' failed for one reason. The truth value of a counterfactual can be changed by the inclusion of 'new' beliefs or exclusion of 'old' beliefs which are not made more or less probable by the assumption that the antecedent A is true.

---

[13] The term 'stock of knowledge or beliefs' appears first, I believe, in Ramsay's 'General Propositions and Causality'. The idea here is that sometimes a connection between antecedent and consequent is relevant and sometimes it is not. Stalnaker, in 'A Theory of Conditionals', uses this as a point of departure from the positivist paradigm.

[14] Exactly how this procedure was to take place varied enormously from theory to theory. Chisholm and Mackie proposed 'thought experiments', Goodman proposed linguistic convention or "entrenchment".

Consider once again, "If I had struck this match it would have lit." This counterfactual can be true only if I had struck the match correctly. But nothing in the assumption "I struck this match" makes it more or less probable that I struck (or can strike) the match correctly. Inclusion of "I struck the match correctly" in the extended antecedent cannot be justified strictly by the assumption that the antecedent is true.[15]

In response positivists attempted to employ a theory of probability. For example, in the case above, all other things being equal and since I am a smoker I probably would have struck the match correctly. But the probability of different occurrences will change as the situation changes. If I had always used a lighter, for example, the probability would change. Or if I had been a different person - a non-smoker, perhaps - the probability would be changed.

Positivists must either assume that the truth of a counterfactual will vary with the context of its assertion or assume a constant 'background' probability.[16] But the latter option leads to incorrect analyses. Consider the following example.

I think it is quite probable that Oswald killed Kennedy, that he was working alone, and that there was no second killer waiting. But I

---

[15] Indeed, it can only be justified by the assumption that the consequent is true - but 'affirming the consequent' is not recognized as a valid rule of inference for material and strict conditionals.

[16] By 'background probability' I mean something like 'prior probability' as defined by Bayes' Theorem. See Skyrms, "Choice and Chance", pp. 156-159, for a formal presentation and Salmon, "Logic", pp. 135-138 for an informal presentation. For a more contemporary presentation see Giere, "Understanding Scientific Reasoning", second edition (it's not in the first), pp. 107-108.

> think it is slightly probable that Oswald was innocent, and that someone else killed Kennedy. I think it is... negligibly probable that Kennedy was not killed at all.[17]

Suppose, then, that we assume that Oswald did not kill Kennedy. Whether or not Oswald killed Kennedy the fact remains that Kennedy is now dead. Therefore, we should assert:

[k1] If Oswald had not killed Kennedy, someone else would have.

and not:

[k2] If Oswald had not killed him, Kennedy would not have been killed.

If the probability that Kennedy is dead remains fixed then we must assert that someone else killed Kennedy. But to allow that the probability can be changed by the context of the assertion that Oswald did not kill Kennedy means two things. First, the truth value of a counterfactual will depend on the context in which it is asserted. Thus the first major thesis of the positivist program fails. And second, the analysis of counterfactuals which follows will no longer be an analysis in terms of the material and strict conditionals. Direct inference from the world in which the counterfactual is asserted will no longer be possible.

The reason for this is as follows. Any two counterfactuals may be asserted in a different context. The 'extended antecedent' for any two counterfactuals may differ. But many of the rules of inference which apply to material and strict conditionals require that the context of assertion remain constant.

---

[17] Lewis, "Counterfactuals", p. 71.

First. Contraposition. If A->B is true then so is -B->-A. But the extended antecedents may be as follows: (A&C)->B and (-B and -C) -> -A. Since C and -C are contradictory, contraposition fails.

Second. Transitivity. If A->B and B->C are true then so is A->C. But the extended antecedents may be as follows: (A&D)->B and (B&-D)->C. Since D and -D are contradictory then transitivity fails.

These rules fail because a third rule fails. The antecedent of a counterfactual cannot be strengthened without possibly changing the truth value of the counterfactual. A->B might be true while (A&C)->B might be false. The failure of this rule was implicit in the use of an extended antecedent. The positivist program was doomed from the start.

3. THE MODEL-MODAL ANALYSIS OF COUNTERFACTUALS

Positivist analyses of counterfactual conditionals failed. The model-modal analysis was developed as a response to this failure. In this section I shall examine the development of the model-modal analysis.

The idea behind the model-modal analysis is this. Counterfactual operators such as 'could' and 'would' ought to be analysed in the same manner as the modal operators 'possibly' and 'necessarily'. The analysis of these modal operators involved a reference to possible worlds. A proposition is necessarily true if it is true in all

possible worlds. A proposition is possibly true if it is true in at least one possible world.

There are two key theses to the model-modal analysis of counterfactuals which distinguish it from standard positivist analyses. They are as follows.

First. Counterfactual propositions are not analysed in isolation but rather are analysed from within the context of a possible world. A technical notion of 'accessibility' is employed which, following Kripke semantics for modal logic, defines the set of worlds which are 'possible' relative to the actual world.

Second. The possible world to be employed is not described by inference from some postulated change in the actual world but is rather selected from a set of possible worlds on the basis of some relevant similarity with the actual world. A formal 'selection function' is employed to select from the set of worlds in which a counterfactual antecedent is true the world which is 'closest' or most similiar to the actual world.

The use of one of these theses entails the use of the other. That the use of the second entails the use of the first is obvious. It is impossible to select a possible world on any basis unless there is first a set of possible worlds from which to select. Demonstrating that the use of the first entails the use of the second more difficult. It is not clear that the use of a possible world entails the use of a selection function.

To demonstrate that a selection function must be employed if possible worlds are employed let us consider some early attempts to employ possible worlds which are constructed or directly inferred from the actual world.

First attempts to employ the modal analysis involved the 'construction' of possible worlds by inference from the actual world. In general, the method was this. Imagine that a counterfactual antecedent A is true. Conjoin A with a 'stock of beliefs' about the actual world. Adjust this stock of beliefs to eliminate any contradictions with A and to accommodate any probable consequences of A. The resultant stock of beliefs, in conjunction with A, is the new possible world. If the consequent B is true in this possible world then the counterfactual A->B is true in the actual world.

These first attempts have much in common with positivist approaches sketched in section 2 and consequently suffer the same problems. The crucial difficulty is this: how is it possible to determine the probable consequences of assuming that a counterfactual antecedent A is true? Any number of probable consequences might follow since the context in which the counterfactual is asserted may vary. Consequently, any number of possible worlds might be constructed. The crucial objection may thus be restated as follows: which possible world, of all those constructed, is the possible world which determines the truth (or falsity) of the counterfactual?

The underlying motivation for this criticism is suggested by the criticism Stalnaker brings against the 'constructivist' approach. The

use of probability, suggests Stalnaker, failed because it was subjective; it measured for belief, not truth. According to Stalnaker,

> It is crucial... that the answer not be restricted to some particular context of belief if it is to be helpful in finding the definition of a conditional function.... the truth function should not in general be dependent on the attitudes which anyone has towards those propositions.[18]

According to Stalnaker the truth of a counterfactual does not depend on someone's psychological state or attitude. Rather, the truth of a counterfactual depends on the actual state of affairs in some world very similiar to our own. This notion of similarity is developed in detail by Stalnaker and in somewhat different detail by David Lewis and forms the second key component of the model-modal theory. "The concept of a possible world is just what we need to make this transition, since a possible world is an ontological analogue of a stock of hypothetical beliefs."[19]

It is important to note that Stalnaker's objection does not apply against construction per se. Rather, the objection is against the assumption that one and only one possible world may be constructed for any given counterfactual. Given that more than one possible world may be constructed for any given counterfactual any selection of a particular possible world will be arbitrary unless mediated by an 'ontological analogue', a set of possible worlds from which a

---

[18] Stalnaker, 'A Theory of Conditionals', p. 102.

[19] Stalnaker, 'A Theory of Conditionals", p. 102.

non-arbitrary selection function selects the closest or most similiar member.

There are two key components to Stalnaker's analysis. First is a technical notion of 'accessibility' which, following Kripke semantics for modal logic, defines what worlds are 'possible' relative to the actual world. Second is a 'selection function' which picks out from that set of accessible possible worlds one world which is the 'closest' or most similiar to the actual world.

The core idea and the core problem for the model-modal analysis is the problem of specifying what counts as accessibility and, within the terms of that notion, what is to count as closeness. In other words, once we have decided to analyze counterfactuals in terms of possible worlds, the core idea and the core problem will be centred around the selection of such worlds.

4. THE SELECTION FUNCTION

To be effective, the model-modal analysis will have to provide criteria for the selection of one particular possible world from a set of possible worlds. The criteria for such a selection, broadly stated, will be based on the similarity between a possible world and the world in which a counterfactual is asserted.

In this section I outline the first attempt to provide such criteria for counterfactuals: Stalnaker's selection function. I proceed as follows. First. I define Stalnaker's terminology in terms

of the model structure defined in chapter one. Second, using this terminology I describe Stalnaker's selection function. Third. I examine Stalnaker's use of the selection function to establish counterfactual semantics. Fourth, I examine the assumptions implicit in Stalnaker's proposal and point to problems of circularity.

Let us adopt some terminology. We shall call the world in which the counterfactual is asserted the 'base world'. The base world is sometimes also referred to as the 'actual world' but for now let us use the ontologically neutral 'base world'.

The base world is a possible world; it is a member of the set K, defined in chapter one, of possible worlds. In particular, the base world is 'G' in the triple (G,K,R).

The possible world which is to be selected as the closest possible world is a member of the set K of possible worlds. In chapter one this selected possible world was denoted using the letter 'H' and described as a 'model' of the base world G. H is a member of K.

H is selected from K by means of a selection function 'f' which takes as arguments a possible world G and a proposition A and has as its value the selected world H. The proposition A is the antecedent of some counterfactual as described in sections 2 and 3 of this chapter.

In chapter one I defined the selection function as a part of the relation R in the triple (G,K,R). Stalnaker and Lewis, the chief architects of the modal analysis of counterfactuals, define the

selection function separately. Stalnaker employs the formal notation:

f(A,a)=b

where 'f' is the selection function, A is the counterfactual antecedent A, 'a' is the base world G, and 'b' is the selected world H. In my own terminology the selection function is a part of the relation R and is expressed:

R(A,G)=H

where it is understood that this is only a partial definition of R.

Having now defined the terminology we may now examine the selection function in more detail.

Stalnaker's selection function has four rules or conditions which are intended to ensure that (i) an appropriate selection is obtained if the antecedent is true in the base world, and (ii) a 'similarity ordering' is obtained if the antecedent is false in the base world.[20]

The four rules are as follows. For any antecedents A and A', for all base worlds G, and for all selected members of K, H:

[1] "A must be true in R(A,G)". This condition ensures that the antecedent is true in the selected world.

[2] "R(A,G)=lambda only if there is no world possible with respect to G in which A is true." In cases where the selection function would select a world which is not a member of K - an 'impossible world' - the notation 'lambda' indicates the selection of an "absurd world".

---

[20] See Stalnaker, 'A Theory of Conditionals', pp. 104-105.

For example, if the antecedent A were "There is a round square" then there is no possible world in which A can be true.[21]

[3] "If A is true in G then R(A,G)=G." This rule requires that the base world be selected as the most similiar world if the antecedent is true in G. This rule allows the analysis to be applicable to all conditional statements and not just counterfactuals.

[4] "If A is true in R(A',G) and A' is true in R(A,G) then R(A,G)=R(A',G)." This rule allows the ordering of possible worlds H on the basis of the similarity with the base world G. Stated in precisely this way, it establishes that if some possible world H is chosen as the closest possible world then no other possible world H' can be chosen which is closer.

The selection function, thus defined, serves to select the closest possible world on the basis of the identity of truth values of propositions in the base world and equivalent propositions in some possible world. Each proposition might be characterized as a 'potential antecedent'; the world selected will be just like the actual world except for differences caused by the assumption that the counterfactual A is true. The world which incorporates the fewest number of changes from the actual world given that A is true will be selected.

---

21 "Possible relative to G" will assume greater importance as we discuss variations possible on the 'accessibility relation' which defines what is 'possible relative to G' below.

Let us now examine how Stalnaker employs the selection function to generate counterfactual semantics.

Stalnaker symbolizes the counterfactual conditional with the "corner conditional" connective. In Stalnaker's notation the counterfactual "If A had happened then B would have happened" is symbolized "A>B".

Truth for the corner conditional is defined in terms of the selection function described in section 3:
A>B is true in G if B is true in R(A,G)
A>B is false in G if B is false in R(A,G)

In some base world G the counterfactual "If A occurred, B would occur" is true if, in some world selected by the selection function R(A,G) in which A is true, B is also true. The same counterfactual is false in G if, in the world selected by R(A,G) in which A is true, B is false.

For example, suppose the counterfactual, "If this match were struck, it would light." Suppose further that the match has not been struck in this world (G). Some world H is selected by the selection function; H is the closest possible world to G in which the match has been struck. If, in H, the match lights, the counterfactual is true. If it sputters and dies, the counterfactual is false.

Stalnaker's corner conditional is distinct from the material or the strict conditional. In particular, laws of inference which apply to the material and the strict conditional do not apply to the corner

conditional. First, strengthening the antecedent may change the truth value of the corner conditional. Second, the corner conditional is not transitive. And third, contraposition is not valid for the corner conditional.[22]

These rules are not valid for the corner conditional because the selection function may select different worlds given different arguments. Suppose, for example, the antecedent of some counterfactual A->B is strengthened to (A&C)->B. Two different sets of arguments result and the values of each use of the selection function may vary. For example, R(A,G)=H while R([A&C],G)=H'. Transitivity and contraposition fail for the same reason.

Stalnaker's semantics are attractive because they appear to solve some long standing problems for positivism. Stalnaker suggests two. First, Hempel's paradox of the black raven is solved since the statement "If x is a raven then it is black" is not equivalent to its contrapositive "If x is not black then x is not a raven." Second, Goodman's "new riddle of induction" is solved because worlds in which the predicate 'blue' is employed are closer to G than worlds in which the predicate 'grue' is employed.[23]

Implicit in Stalnaker's proposal is an assumption of natural systemic coherence. Let us now examine this assumption.

---

[22] See Stalnaker, 'A Theory of Conditionals', pp. 106-107.

[23] Whether these are satisfactory solutions is, I think a question still to be debated.

The assumption is essentially as follows. The closest possible world will be some world in which laws of nature and other regularities will be similiar to those of the base world regardless of what particular facts will have changed.

It is through such implicit criteria that Stalnaker asserts that possible worlds may be investigated empirically. Tendencies and laws of nature will remain the same in the selected world as in the actual world unless explicitly stated otherwise. for example, Stalnaker writes:

> When I wonder, for example, what would have happened if I had asked my boss for a raise yesterday, I am wondering about a possible world that I have already roughly picked out. It has the same history, up to yesterday, as the actual world, the same boss with the same dispositions and habits.[24]

The selection of some possible world H is therefore going to require at least a partial description of the base world G. Since the selection function is such a crucial component of the modal analysis of counterfactuals any difficulty in the selection function will be a difficulty for the entire theory. But we can see such a difficulty beginning to develop already.

The base world must be described in order to specify some possible world H which is closest to the base world. But the description of the base world will require that some counterfactual statements, such

---

[24] Stalnaker, 'A Theory of Conditionals', p. 112.

as 'my boss's dispositions', be known to be true.[25] To know that such counterfactuals are true will in some cases require reference to a further possible world H'. A revised description of G will be required. This will require more true counterfactuals. An infinite regress is generated.

5. THE ACCESSIBILITY RELATION

In this section I shall examine the use of an 'accessibility relation' in the possible worlds analysis of counterfactuals.

I will suggest in this section that the accessibility relation serves the same function as the selection function outlined in section 4 while apparently avoiding some of the difficulties mentioned immediately above.

This section is structured as follows. First I describe the accessibility relation as employed by Kripke for modal logic and demonstrate how it fits into the model structure described in chapter one. I then explain the use of the accessibility relation by Stalnaker and Lewis in the possible worlds analysis of counterfactuals. Finally, I show how the accessibility relation, like Stalnaker's selection function, is used to generate counterfactual semantics.

---

[25] See Goodman, "Fact, Fiction and Forecast", chapter 2, section 3, for a discussion of why dispositions are counterfactuals.

I shall now explain how the accessibility relation is employed in Kripke semantics.

By 'accessible' to some world G we mean 'possible relative to G'. Kripke, who first defined the accessibility relation as part of the triple (G,K,R), describes it thus: "If H1 and H2 are two worlds, H1RH2 means intuitively that H2 is possible relative to H1, i.e., that every proposition true in H2 is possible in H1."[26]

The set of worlds K which are possible relative to G is defined by R prior to the use of some member of K in the analysis of some counterfactual. R may be defined in any number of ways; there is only one a priori restriction: that it be reflexive. "Every world H is possible relative to itself, since every proposition true in H is, a fortiori, possible in H."[27] Notice that this is the only restriction applied to membership in K in my own model structure described in chapter one.

Kripke did not intend R to be just one particular relation; by varying R we can vary the nature and extent of the set K. For example, if R is reflexive only we get the M-model structure. If it is reflexive and symmetrical we get the Brouwer system; reflexive and transitive, S-4; and if reflexive, transitive and symmetrical, S-5.[28]

---

[26] Kripke, 'Semantical Considerations on Modal Logic', p. 64.

[27] Ibid.

[28] See Kripke, p. 64. The modal systems named are defined by the addition of different modal axioms; see Loux pp. 16-17 (among other sources).

There is much more which could be said about the use of R in possible world semantics but enough has been said for the purposes of this paper. Let me now examine how it is employed in the possible worlds analysis of counterfactuals.

In the possible worlds analysis of counterfactuals the accessibility relation R has three major functions. First, it defines some set of possible worlds K which are possible relative to some base world G. Second, it eliminates some worlds which are not possible relative to G (this is the analogue of Stalnaker's 'absurd world' described above). And third, R orders the set of possible worlds K according to their 'relative similarity' to the base world G.

Stalnaker has little to say about the accessibility relation. He uses it primarily to define the set K and eliminate the 'absurd world'. Stalnaker's use most clearly follows Kripke's definition and like Kripke he acknowledges that different definitions of R (as transitive or symmetrical) may be employed. He does not appear to have recognized that R may be used to order possible worlds by comparative similarity to some base world.[29]

---

[29] It might be tempting here to say that Stalnaker states four conditions whose function and effect is in fact simply to define the set K and eliminate the absurd world. This would be slightly inaccurate. On my own definition of R I could say that. But Stalnaker distinguishes between the accessibility relation and the selection function. I agree that in fact Stalnaker's four conditions simply define K and eliminate the absurd world but I think Stalnaker wants to hold that the selection function is distinct from the accessibility relation - otherwise, why introduce it? But I think we can see why I want to merge all of this (selection, ordering, salient qualities) into one relation R even though few (if any) of the model-modal theorists want to.

Lewis, in "Counterfactuals", defines accessibility more precisely. He has developed a system of diagrams which assign to each possible world a "sphere of accessibility" around that world. In Lewis's terms the world in question is called 'i' and the sphere is called $i. The equivalent to 'i' in this paper is some member of K; that is, G. Inside the sphere are 'possible worlds'; that is, worlds which are 'possible' relative to i.

Lewis first adapts his sphere of accessibility to the Kripke semantics from which it is drawn. A proposition 'A' is necessary in 'i' (or G) if and only if, for each world in the sphere of accessibility $i around 'i', A is true. A proposition 'A' is possible in 'i' (or G) if and only if, for some world in $i, A is true.

As necessity and possibility are determined by the values of one proposition, A, in the worlds in $i, conditional necessity and possibility are determined by the values of two propositions, A and B, in the worlds of $i. A strict conditional for some i is true if, for every world in which A is true, B is also true. A 'variably strict conditional' for some i is true if, for some worlds (but not all) in which A is true, B is true. Counterfactuals are variably strict conditionals; A is not true in the base world and both A and B are true in at least one world in $i.

Lewis's system of spheres $i is the set of possible worlds K which is defined by R relative to some base world G (or, in Lewis's terms, i). The structure of K - that is, the precise ordering of the worlds in K with respect to G - will be determined by R. Some worlds may be

closer to the centre of the sphere (G) than others. Exactly what the ordering of the set of possible worlds K (in $i) will be determined by exactly how we specify R.

For Lewis, worlds are ordered in $i according to their similarity with G (or i) by R. Worlds may be more or less similiar to G according to various criteria and consequently more or less accessible. The world which is the most accessible is the most similiar; the world which is the most similiar is the most accessible. There may thus be a system of spheres around G such that, the wider the sphere, the less similiar and the less accessible. Such a system of spheres has the following properties: it is centred on G and it is nested, thus establishing an ordering of similarity to G.

For any G, since there may be more than one accessibility relation R, there may be more than one sphere of accessibility around G. Different spheres of accessibility will be defined by different definitions of the accessibility relation R. We have seen above that R could be either transitive or symmetrical and transitive. More precise specifications are possible. For example, R might define the set K such that only worlds in which the laws of nature are the same as in G can be members of K. Another example: R might define the set K such that only worlds in which there are blue cows may be members of K.

So defined, the accessibility relation may replace the selection function. Like the selection function, the accessibility relation determines counterfactual semantics. Let us examine how this is done.

Lewis, like Stalnaker, employs the closest possible world in which some counterfactual antecedent is true to determine whether the counterfactual is true in some base world. There is no need to employ a selection function to determine which world is closest since the accessibility relation R orders members of $i on the basis of relative similarity.

The world selected will be the closest possible world in some sphere $i as defined by some R. Accessibility may be restricted to include only those worlds in which the antecedent A is true.[30] Selection is automatic given some R and some antecedent A.

Truth in Lewis's semantics is defined for the most part in exactly the same manner as Stalnaker's. The counterfactual statement "If A had happened, B would have happened" is true if and only if B is true in the closest world H in some sphere of accessibility $i around G.

Consider the following example:

If Allan had come to the party it would have been a good party, but if Allan and Bill had come to the party they would have fought and it would have been a bad party, but if Allan and Bill and Charlie had come it would have been peaceful and thus a good party.

The statements are symbolized thus:

A (box-arrow) G

(A&B) (box-arrow) -G

(A&B&C) (box-arrow) G

Each of the three statements refers to a sphere $i which is less similiar to G than the one previous. In the closest world in which

---

[30] See Lewis, "Counterfactuals", p. 12.

only A is true G is also true. But in the closest world in which both A and B are true G is not true. It is possible for the conclusion (that it was a good party) to be different in each successive sphere of possible worlds.

Lewis's semantics may be seen as a refinement of Stalnaker's. Unlike Stalnaker, who uses only one counterfactual operator, Lewis employs two. First, the 'box-arrow' symbolizes the counterfactual operator 'would'. Second, the 'diamond-arrow' symbolizes the counterfactual operator 'could'.

Lewis's semantics differs from Stalnaker's because of its use of the 'could' operator. The 'could' operator is employed in cases where there is a tie; two worlds are equally close to G and in one B is true and in the other B is false. The counterfactual statement "If A had happened, B could have happened" is therefore true if and only if, for some closest world H in $i around G, B is true.[31]

In the next section I will explain why Lewis permits such ties and why Stalnaker does not. I will then explain how Lewis proceeds to break such ties using 'salient qualities' which are determined by the context in which a counterfactual is asserted. As we shall see, the use of such salient qualities leads Lewis into the same difficulties of circularity and regress which plagued Stalnaker and which plague all versions of the model-modal theory.

---

[31] See Lewis, "Counterfactuals", pp. 16 and 22.

6. CONTEXT AND SALIENCE

While both Lewis and Stalnaker believe possible worlds may be more or less similiar to some base world Lewis believes that two worlds may be of (sufficiently approximate) equal similarity while Stalnaker asserts that for each possible world H any other possible world H' is either more or less similiar to G than H.[32]

That amounts to the following. For Stalnaker, the selection of some H of K may be determined strictly by the use of the selection function. For Lewis the selection function is replaced by the accessibility relation and more than one world may be selected; the selection of some particular H which is closest to some G might also require the use of a more precisely defined accessibility relation.

As mentioned above, both Stalnaker and Lewis use possible worlds as models H1, H2,... which are related by similarity to some base world or 'original' world G. Accordingly we may characterize the distinction between Stalnaker and Lewis as nothing more than a difference in the way each defines R.

Stalnaker and Lewis restrict K differently. Stalnaker defines R only in terms of logical possibility relative to G (thus following Kripke most closely) while Lewis restricts R in terms of logical

---

[32] Lewis expresses this distinction in terms of a "conditional excluded middle" which is: "Either 'If A occurred B would occur' is true or it is false." According to Lewis, Stalnaker's system, which employs the conditional excluded middle, is a "special case" of his. ("Counterfactuals", p. 78.) Van Fraassen, by contrast, characterizes Lewis's system as a supervaluation of Stalnaker's (In 'Hidden Variables in Conditional Logic').

possibility as well as some salient qualities relative to G. Either definition of R is allowable in the model structure defined in chapter one.

Different accessibility relations may be employed at different times. These accessibility relations are defined by Lewis in terms of salient qualities. Which qualities count as salient qualities is determined by the context in which a counterfactual is asserted.

Let us examine precisely how that works.

Consider an analysis of some counterfactual conditional in which two distinct accessibility relations are possible such that two different worlds are selected, each of which is as close to the base world G as the other. Such a counterfactual is 'ambiguous'; there is no obvious means of selecting which of the two worlds should be chosen.

An example of this type of situation may be found in Quine's "Word and Object".[33] Consider the following pair of statements:

If Caesar were in command (in Korea) he would have used the atom bomb.

If Caesar were in command he would have used the catapults.

Such pairs caused endless problems for positivists. Quine writes,

> What traits of the real world to suppose preserved in the feigned world of the contrary-to-fact antecedent can be guessed only by spinning from a sympathetic sense of the fabulist's likely purpose in spinning his fable.[34]

---

[33] P. 222.

[34] Ibid.

Unless some non-arbitrary method is employed such pairs must remain unresolved.

In Lewis's analysis of counterfactuals such pairs are resolved with reference to salient qualities. Such salient qualities are defined by the context in which the counterfactual is asserted. In the example above, if the salient quality was "Caesar's primitive technology" then the second counterfactual would be true. If it were "Caesar's ruthlessness" then the first would be true.

Lewis selects salient qualities from a context in which a counterfactual is asserted. They are those qualities which are relevant to the counterfactual in question. The occurrence of salience depends on the existence of some person to whom some qualities may be more important than others.

> When i (in some $i) is a thing with a point of view - say a person or an animal - then some things are more salient than others from the point of view of i. They loom larger in his mental life; they are more important to him; they come more readily to the centre of his attention.[35]

Salient qualities function to determine an ordering of the possible worlds in the sphere $i.

> A sphere around i is to be any set of things in the ken of i such that all those in the set are more salient to i than any of those outside. Assuming that comparative salience orderings have the definite properties of weak orderings, then those sets do indeed comprise a system of spheres $i.[36]

---

[35] Lewis, "Counterfactuals", p. 113.

[36] Ibid.

In the previous section I explained how the accessibility relation replaces Stalnaker's selection function. At that time it seemed that some of the difficulties inherent in Stalnaker's proposal might be avoided. Now let me show how these difficulties return with the use of context and salience to define the accessibility relation.

The close relation between salience and the concepts of context dependence and one's mental life shows in turn the close dependence between context dependence and the use of model structures. Recall from above that the accessibility relation R may vary considerably. If the selection of some accessibility relation R is determined by one's point of view (or perhaps one's world view) then the determination of which of two possible worlds H1 and H2 is most accessible (closest) to G depends directly on some description or opinion about G.

But now exactly the same problem occurs. A description of G or opinion about G might require the assertion that some counterfactual statement is true. But if such a counterfactual description is required then reference to some other possible world is required. This in turn requires another description. The regress continues.

Notice that Stalnaker was shown to face exactly the same problem in section 4. If we need to describe G to select H and we need to select H to describe G then either vicious circularity or vicious infinite regress occurs.

In chapter one, section 11, I explained this problem strictly in terms of the model structure defined above. Stalnaker and Lewis both employ forms of this model structure. It should therefore not be suprising that both face the same problem.

Once again I want to emphasize the fact that the core idea of the possible worlds analysis and its core problem are one and the same: the selection of some possible world H which is closest to some base world G.

Chapter Three

POSSIBLE WORLDS AND THE EVALUATION OF CAUSAL PROPOSITIONS

1. INTRODUCTION

In this chapter I shall examine an example of how the possible worlds analysis of counterfactuals is used and show that in such a particular instance it collapses into vicious regress and circularity.

The objection employed will be of precisely the form suggested in chapter one, section 11 and chapter two, sections 4 and 6.

The example selected is the problem of evaluating causal propositions. By the 'evaluation' of some proposition I mean something distinct from the 'analysis' of some proposition.

To provide an analysis is to specify the truth conditions for some class of propositions. To provide an evaluation is to specify the truth values of some particular proposition. Obviously the analysis and evaluation of causal propositions are heavily related and I will draw heavily on the analysis described in chapter two.

To evaluate a proposition can mean more than just the specification of truth values. Other important values may be specified as well. For example, in this chapter, part of the problem of causation is the problem of determining whether a true proposition is a causal proposition.

I shall consider this latter problem as well.

This chapter is structured as follows. The problem of causation is defined using examples and the failure of positivist attempts at a solution is outlined. The possible worlds solution is detailed and shown to stand against major objections. The main objection, which corresponds exactly to that outlined in chapters one and two above, is applied and shown to succeed.

A more precise structure is as follows.

In section 2 I distinguish problems of the analysis of causal propositions from problems of evaluation. I use two examples to precisely define the latter, the first of which defines the problem of determining whether a causal proposition is true, the second of which focusses on the problem of determining whether a particular proposition is a causal proposition.

In section 3 I outline the positivist approach to the problem of evaluation. I show that positivism fails because no method of confirmation could be defined which solves problems posed by the examples given in section 2. A dilemma is posed: if confirmation is attempted directly from observed regularities causal propositions

cannot be distinguished from other regularities; yet if confirmation is attempted counterfactually then causal propositions are confirmed in the absence of any regulartities.

In section 4 I introduce the possible worlds theory as a solution to the problem of evaluation. First I show why the possible worlds theory can confirm regularities counterfactually while positivism failed. Second I show the precise counterfactual form to which a causal proposition must correspond. And third I show how this form establishes causal dependence, thus distinguishing between true causal propositions and other propositions.

In section 5 I outline major objections to the possible worlds solution and show why they fail. I sort these objections into four categories: (i) objections which assert that causal dependency is empirically established, (ii) objections which assert that it is ontologically established, (iii) objections which assert that the wrong truth values are determined for particular propositions, and (iv) objections which assert that no truth values are determined. The possible worlds solution is shown to hold against all four categories.

In section 6 I detail the main objection to the possible worlds solution. First I reconsider the two problems from section 2. Second I provide the procedure for solving these problems using the method outlined in section 4. Third I show that exactly the same objection as that outlined in chapter one, section 11, succeeds against the possible worlds solution. And fourth, I show in the light of that successful objection why previous objections failed.

In section 7 I suggest that the world view analysis described in chapter one might be considered the only response to the main objection. In addition I motivate the world view analysis's prime thesis that observation is theory dependent in terms of the main objection.

2. THE PROBLEM(S) OF CAUSATION

In this section I outline for the first time the problem of causation.

In fact there are a number of problems. I sort them roughly into two categories: problems of analysis and problems of evaluation.

I summarize both the positivist and the model-modal approach to each of these categories and restrict discussion in this chapter to problems of evaluation only.

There are two problems of evaluation. First, how do we infer that a particular causal proposition is true? And second, how do we infer that a particular proposition is a causal proposition? I define each of these problems more precisely using examples.

There are a number of ways we may express the "problem of causation":

1. How do we justify the inference from the existence of a 'cause' to the existence of an 'effect'?

2. How do we infer from the existence of a 'cause' to the existence of a particular 'effect.'

These two statements of the problem correspond roughly to Hume's 'problem of induction' and Goodman's 'new riddle of induction' respectively.

3. How do we infer that a particular causal proposition O(C->E) is true?

4. How do we infer that a particular proposition O(A->B) is a causal proposition?

These two statements of the problem correspond roughly to the 'problem of backward causation' and the 'problem of epiphenomena and accidental generalizations' respectively.

5. How do we analyze causal propositions?

6. How do we evaluate causal propositions?

These two statements of the problem correspond roughly to statements (1) and (2) and statements (3) and (4) immediately above respectively.

We take as a starting point the following solutions to problems (5) and (6) respectively:

(A) Causal propositions are counterfactuals and hence are analysed as counterfactuals.

(B) A true causal proposition is a counterfactual in which the consequent (the effect) 'depends' on the antecedent (the cause).

We may now state the problem of causation as follows:

7. How do we analyze counterfactuals?

8. How do we establish that, for a particular counterfactual, the 'effect' depends on the 'cause'?

In chapter two we considered problem (7). In particular we distinguished between the positivist answer to problem (7) and the model-modal answer as follows:

(C) The positivist response: (i) counterfactuals are analysed independently of the context in which they are asserted, and (ii) are analysed as variations of material or strict conditionals such that true counterfactual statements are inferred directly from evidence in the world in which they are asserted.

(D) The model-modal response: counterfactuals are analysed from within the context of a possible world determined by the context in which the counterfactual is asserted, and (ii) are analysed indirectly on the basis of the similarity of the possible world to the world in which the counterfactual is asserted.

In this chapter we shall consider problem (8). To begin we distinguish between the positivist answer to problem (8) and the model-modal answer as follows:

(E) Dependence on a cause by an effect is inferred directly from the world in which the dependence is asserted from previously established 'regularities' or 'constant conjunctions' of cause and effect.

(F) Dependence on a cause by an effect is inferred indirectly from selected possible worlds such that in the closest possible world in which there is a particular cause there is a particular effect and in the closest world in which that particular cause does not exist that particular effect does not exist.[37]

I have identified two problems of evaluation:

1. How do we infer that a particular causal proposition is true?

2. How do we infer that a particular proposition is a causal proposition?

I shall now define each of these problems more precisely with an example for each of the problems stated above.

Problem 1. Consider the causal proposition [s1] "Striking a match causes it to light."

We may express this counterfactually as: [s2] "If match m had been scratched it would have lighted."

---

[37] Lewis, in 'Causation', defines the alternative approaches in just these terms citing as his source two definitions offered by Hume in his "Treatise".

Our knowledge of matches, chemical processes and relevant conditions leads us to affirm that this counterfactual is true.

But on equal grounds we may assert that the following counterfactual is true: [s3] "If match m had been scratched it would not have been dry."

Here is the reason. Match m did not light. Since m did not light (and assuming that it was well made, scratched correctly, etc.) and since dry matches light then m could not have been dry.

As the counterfactual [s2] corresponds to the causal proposition [s1] so does [s3] correspond to [s4]: "Striking a match causes it to become wet."

The question is this. Given that equal evidence supports both cases, why do we suppose that [s1] and [s2] are true and [s3] and [s4] are false?[38]

Problem 2. Suppose that a kettle of water is heated. Bubbles form and steam rises as the temperature of the water increases.

We assert quite naturally [s5] "Heating the kettle caused the bubbles to form," and [s6] "Heating the kettle caused the steam to rise."

---

[38] This example has been adapted from Goodman, "Fact, Fiction and Forecast", p. 14.

We can express these counterfactually. Corresponding to [s5] is [s7]: "If the kettle were heated bubbles would form." Corresponding to [s6] is [s8]: "If the kettle were heated steam would rise."

Equally true is the following counterfactual: [s9] "If steam rises then bubbles form". But we do not consider the corresponding [s10] to be a true causal statement: "Steam rising caused the bubbles to form."

So the problem is this. All three counterfactuals [s7], [s8], and [s9] are true and yet only the counterfactuals [s7] and [s8] correspond to causal propositions. What distinguishes [s9] from the preceding?

In the evaluation of causal propositions we have two problems. First, why do we assert that some causal propositions and their corresponding propositions are true while asserting that others are false? And second, why do we assert that some true counterfactuals correspond to true causal propositions while others, equally true, do not?

In both the positivist and the model-modal tradition solutions to these problems have rested on the following intuitions: first, a proposition is a causal proposition if the 'effect' depends on the 'cause'; and second, a causal proposition is a true causal proposition if, for the same 'cause' and the same 'effect', the cause does not depend on the effect.

But the two traditions split when we ask what sort of intuition this is and when we ask how this intuition is to be given stronger epistemological support. To positivists, the intuition was one about the nature of observed regularities in the world and could be substantiated with further empirical evidence. To model-modal theorists the intuition was one about the nature of certain counterfactuals and could be substantiated by a logical solution to the problem of counterfactuals.

3. THE POSITIVIST RESPONSE AND WHY IT FAILS

Positivists attempted to evaluate causal propositions by direct inference from previously established regularities in the world.

In this section I shall survey positivist attempts to confirm such regularities such that a dependence of an effect on a cause could be inferred.

I first outline attempts to confirm causal regularities directly from observed instances of such regularities. The focus is primarily on the attempt to determine necessary and sufficient conditions which hold between a cause and an effect. I show that such attempts fail to establish causal dependency.

I then outline attempts to confirm causal regularities counterfactually from the observed absence of a cause given the absence of an effect. I show that such attempts allow the

confirmation of causal regularities in the absence of any observation of such a regularity.

In effect I propose a dilemma for the positivists. On one hand attempts to confirm causal regularities on the basis of observed instances only do not confirm regularities such that the effect depends on the cause while on the other hand attempts to confirm causal regularities on the basis of the lack of counterexamples allow confirmation of regularities which possibly do not exist.

Positivists proposed that causal propositions, which could be expressed as counterfactuals, could be inferred directly from observed regularities, which could be expressed as strict conditionals. The inference was this: from [c1] "in all cases in which A has occurred, B has occurred" to [c2] "if A were to occur, B would occur."

This positivist response splits into two parts: first, an analysis of causal propositions which permits the inference from true material or strict conditionals such as [c1] to true counterfactuals such as [c2]; and second, an evaluation of causal propositions such that statements such as [c1] are known (or believed) to be true.

As we have seen in chapter two the positivist analysis of counterfactuals in terms of material or strict conditionals failed. The history of positivist attempts at such analyses of causal propositions parallels the history of attempts to analyze counterfactuals.

Efforts to form an evaluation of causal propositions in terms of true statements such as [c1] foundered on the problem of confirmation. No statement of the form "in all cases in which A has occurred, B has occurred" could be confirmed which satisfied our intuitions about causal dependency.

Statements strictly of the form (A->B) were too easily disconfirmed. For example, the statement "in all cases where a match is struck, it lights" is disconfirmed by the observation that wet matches do not light.

Concurrent with the analysis of counterfactuals attempts were made to confirm statements such as [c1] such that "relevant conditions" were included. Such statements had an 'extended antecedent' as described in chapter two. Such analyses focussed on two different approaches.

First, "C is a cause of E if and only if C is ceteris paribus sufficient for E."

Second, "C is a cause of E if and only if C is ceteris paribus necessary for E."[39]

The logical notation for such sentences in strict conditional form may be expressed respectively:

---

[39] These sentences are adapted from Sosa, p. 1. Sentences of the first form are proposed by Braithwaite, Hempel and Popper, among others. Sentences of the second form are proposed by Nagel and Raymond Martin, among others.

First: C & (S1 & S2 & ...) -> E

Second: E & (S1 & S2 & ...) -> C

As discussed in chapter two statements of this form failed as an analysis of counterfactuals. Statements of this form as well do not establish causal dependency required for the evaluation of causal propositions.

To demonstrate this let us consider the following example. [s11] "The position of a table top relative to the floor is caused by the length of the legs." Let us assume that [s11] is possibly a true causal proposition.

The sentence [s11] may be expressed using the first or second forms of the notation above. Let us use the strongest form possible, a conjunction of necessary and sufficient conditions:

[s12] "The length of the legs is both necessary and sufficient for the position of the table top relative to the floor."

But the following statement is also true:

[s13] "The position of the table top relative to the floor is both necessary and sufficient for the length of the legs."

The same set of necessary and sufficient conditions which establishes that the effect depends on the cause can be used to establish that the cause depends on the effect. Statements of proposed causal propositions in terms of necessary and sufficient

conditions therefore do not establish the required dependency between cause and effect.[40]

Since statements of the strict conditional form were not sufficient to establish causal dependency some positivists felt that counterfactuals should be employed to establish that the absence of an effect entails the absence of its corresponding cause.

Accordingly we get statements like:

[c3] C causes E iff C & (S1 & S2 & ...) -> E and -E & (S1 & S2 & ...) -> -C

[c4] C causes E iff E & (S1 & S2 & ...) -> C and -C & (S1 & S2 & ...) -> -E

Or the conjunction of [c3] and [c4].

But now the absence of the cause and the effect provides substantive evidence for the assertion of dependency between cause and effect. Hempel's famous paradox of the black raven provides a paradigm example of this sort of difficulty. Consider the statement:

---

[40] This example is from Sosa, p. 3. Sosa cites further examples from Taylor, "Action and Purpose", pp. 35-7. Variations on this basic positivist approach proliferated. Mackie proposed 'INUS' conditions: C is a cause of E if and only if C is an insufficient but necessary part of a condition which is itself unnecessary but exclusively sufficient for E. Davidson and von Wright attempted analyses in which causal dependency is a "primitive" such that the cause "brings about" the effect. None of these variations successfully expressed causal dependency in terms of strict conditionals. See Sosa, pp. 3-8 for a survey of these attempts. primary references are Mackie, 'Causes and Conditions', p. 245, Davidson, 'Causal Relations' in Sosa, pp. 82-95, and von Wright, "Explanation and Understanding", p. 70.

[s14] All ravens are black.

This is confirmed by the inductive generalization:

[s15] In every case in which some entity has been a raven, that entity has been black.

It is also confirmed by the counterfactual:

[s16] In every case in which some entity has not been black it has not been a raven.

Which allows the unusual possibility of being able to confirm that "all ravens are black" without ever having seen a raven.

The positivist is thus faced with a dilemma. On the one hand, if a causal proposition is confirmed using only available evidence (of the form C->E and variants) then causal dependency cannot be established.[41] On the other hand, if a causal proposition is confirmed using counterfactual evidence (of the form -E -> -C and variants) then the causal proposition is confirmed by what is in effect no evidence.

---

[41] Even if, as Goodman suggests, all the "stated" evidence is employed ("Fact Fiction and Forecast", p. 71) since that amounts to no more than an extended antecedent.

4. THE NEW THEORY: CAUSAL DEPENDENCE AND POSSIBLE WORLDS

Model-modal theorists begin with the intuition that the use of the counterfactual (-E -> -C) in conjunction with (C -> E) was a step in the right direction.

Positivists considered the confirmation of causal regularities to be an empirical problem and consequently the confirmation of such a regularity counterfactually posed an empirical dilemma. Model-modal theorists, by contrast, approach the problem as a logical problem.[42]

On the model-modal theory one does not confirm that a particular regularity is a causal regularity by observation and confirmation. Rather, all regularities are confirmed counterfactually and causal regularities are those which are of a particular counterfactual form.

So in the explication of the new theory we have three questions. First, how is it possible to confirm regularities counterfactually in the light of Hempel's paradox? Second, what precisely is the 'particular counterfactual form' to which a true causal proposition must conform? And third, how does this particular form distinguish between true causal propositions and other propositions?

Let us consider the first question first.

Goodman recognized but did not see the importance of the reason for Hempel's paradox.

The statement that a given object, say this piece of paper, is neither black nor a raven confirms the hypothesis that all non-black things

---

[42] I owe this observation to John Baker.

are non-ravens. But this hypothesis is logically equivalent to the hypothesis that all ravens are black.[43]

If instances of the form A->B confirm the proposition O(A->B) then instances of the form -B -> -A equally confirm O(A->B) since the two are logically equivalent. More precisely, the second is the contrapositive of the first.

On the modal analysis of counterfactuals contrapositives are not logically equivalent. Instances of A->B are logically distinct from instances of -B -> -A. The paradox cannot arise if instances of both are required for confirmation since an instance of one is not logically equivalent to an instance of another. In other words, the sighting of a black raven is logically distinct from the sighting of a non-black non-raven and instances of each are required to confirm that all ravens are black.[44]

Positivist evaluations of causal propositions failed because of a flawed analysis of counterfactuals. On the model-modal theory regularities may be confirmed counterfactually without paradox. Consequently causal regularities may be confirmed counterfactually without paradox.

We now turn to the second question. Exactly what is the 'particular counterfactual form' to which a true causal proposition must conform?

---

[43] Goodman, "Fact, Fiction and Forecast", p. 70.

[44] See Stalnaker, 'A Theory of Conditionals'.

I shall begin with an oversimplified definition which will express the main idea.[45]

A true causal proposition O(C->E) is one in which the corresponding counterfactuals C->E and -C -> -E are true. Given the modal analysis of counterfactuals as described in chapter two we can reformulate the definition of a causal proposition as follows:

A proposition O(C->E) is a 'true causal proposition' in some world G if and only if, in the closest possible world to G in which C is true, E is true, and in the closest possible world in which -C is true, -E is true.[46]

As I mentioned, this is a simplified definition although it expresses the main idea. Lewis provides a more complex definition which requires a clear map of the logical territory.

A subset of the set of all propositions is the set of counterfactual propositions. These are expressed O(A->B) and are intended to correspond to a set of counterfactual instances, A1->B1,

---

[45] Lewis employs the simple definition in most of his uses of causal propositions. In particular, see Lewis, 'Causation' and 'Counterfactual Dependence and Time's Arrow'. The statement employed here is from "On The Plurality of Worlds", p. 23.

[46] In most accounts of Lewis's definition (cited immediately above) the definition explicitly mentions only the -C -> -E component. The assumption is that an occurrences of C and E have occurred in G and so G is selected as the closest possible world to itself in which C is true. See Lewis, 'Causation', p. 563: "The dependence consists in the truth of two counterfactuals O(c) -> O(e) and -O(c) -> -O(e)." And in "On The Plurality of Worlds", p. 23, he writes, "Suppose that two wholly distinct events occur, C and E; and if C had not occurred, E would not have occurred either."

A2->B2, ..., An->Bn. A subset of the set of counterfactual propositions are those which occur in pairs such that both O(A->B) and O(B->A) are true. Let us call these 'regularities'.[47]

A subset of such regularities is the set of counterfactual dependencies. Since the items in this subset are regularities and, as such, by hypothesis, true, both O(A->B) and O(B->A) are true. As well, O(-A -> -B) is true while O(-B -> -A) is false. In such a case B would be said to be 'counterfactually dependent' on A. These latter two propositions correspond with sets of instances in just the same manner as the former.

A subset of the set of counterfactual dependencies is the set of causal dependencies. Lewis defines what he calls a "family" of related causes C1, C2, ..., Cn and a "family" of related effects E1, E2, ..., En such that the proposition O(C->E) is a true causal proposition if each pair in the respective families of causes and effects are counterfactual dependencies.

More formally, O(C->E) is a true causal counterfactual if and only if, for each C1, C2, ..., Cn and for each E1, E2, ..., En, the corresponding propositions O(C1->E1), O(C2->E2), ..., O(Cn->En) are instances of counterfactual dependence; that is, for each proposition instances of the counterfactuals Cn->En, En->Cn and -Cn -> -En are true and (at least some) instances of -En -> -Cn are false.

---

[47] This seems to be in accord with normal usage of the term.

We now address the third question. How does this particular form distinguish between true causal propositions and other propositions?

As when I answered the second question, I shall present a simplified version first. Subsequently I shall introduce the corresponding complexity and explain why this complexity is introduced.

If some proposition O(A->B) is a true causal proposition then both the counterfactuals A->B and B->A will be true. This occurs because both A and B are true in the base world G.

But if the proposition O(A->B) is a true causal proposition the counterfactual -A -> -B will be true but the corresponding counterfactual -B -> -A will be false. That is, in the closest possible world in which the cause does not exist, the effect will not exist, but in the closest possible world in which the effect does not exist, the cause will still exist.

We now introduce the complexity. As in the case of the second question, the complexity involves the definition of causal propositions in terms of families of instances.

Structuring the definition of true causal propositions such that it corresponds to families of instances allows Lewis to distinguish between the possible worlds analysis of causation and previous 'regularity' analyses, described above.

In the latter, Lewis argues, dependence is expressed only as a one-to-one dependence between each cause C1, C2, ..., Cn and each effect E1, E2, ..., En. He calls this type of dependence "nomic dependence". By contrast, counterfactual dependence is expressed as a many-to-one dependence such that each cause C1, C2, ..., Cn may cause any of the effects E.

When we assert O(C->E) using nomic dependence we assert {O(C1->E1), O(C2->E2), ..., O(Cn->En)} if O(C->E) is a causal counterfactual. When we assert O(C->E) using counterfactual dependence we assert {O(C1->E), O(C2->E), ..., O(Cn->E) if O(C->E) is a causal counterfactual.

Let us consider an example of this distinction.

Suppose that the air pressure A causes the barometer reading B. The causal counterfactual would then be O(A->B). Different air pressures A1, A2, ..., An cause different barometer readings B1, B2, ..., Bn.

This dependency, expressed as a nomic dependency, is expressed as:

[d1] For each barometer reading, there is a corresponding air pressure on which it depends.

As a counterfactual dependency it is expressed as:

[d2] For each barometer reading, there is some air pressure on which it depends.

If, as in [d1], we assert that there is a particular air pressure on which each barometer reading depends then it becomes true that, if there is a certain barometer reading Bn, there must be a particular air pressure, An. This allows without refutation the suggestion that the barometer reading Bn caused the air pressure An.[48]

If, as in [d2], we assert only that there is some air pressure on which each barometer reading depends then given a particular barometer reading Bn it does not necessarily follow that there is a given air pressure An.

There is therefore an asymmetry in the causal relation. For each pair of counterfactual propositions which establish some regularity, O(A->B) and O(B->A), only one of the corresponding counterfactuals O(-A -> -B) or O(-B -> -A) will be true in all cases. Again, this is a slight oversimplification.

Causal asymmetry is substantiated using the possible worlds analysis of counterfactuals. The counterfactual -A -> -B is true in some world G if in the selected possible world H in which -A is true, -B is true. The counterfactual -B -> -A is analysed in the same manner except that a different possible world, H', is selected.

Since causal propositions are established using sets of instances we may define a set of possible worlds such that some pair of possible worlds corresponds to each instance: (H1,H'1),(H2,H'2), ..., (Hn,H'n).

---

[48] Since -B -> -A would be true.

Almost all instances of the counterfactual -An -> -Bn will be true; that is, in almost all of the corresponding possible worlds Hn in which -An is true, -Bn will also be true.[49]

But by contrast, much fewer instances of the counterfactual -Bn -> -An will be true; that is, in much fewer of the possible worlds H'n in which -Bn is true will -An also be true.[50]

The first counterfactual O(-A -> -B) is therefore stronger than the second counterfactual O(-B -> -A). Consequently the regularity expressed by the corresponding propositions O(A->B) and O(B->A) is asymmetric. A true causal proposition is the proposition O(A->B) which corresponds to the proposition in such a pair which is the 'stronger' of the two.

5. OBJECTIONS THAT FAIL

In this section I want to survey some objections to Lewis's proposal which fail. It will be my intent in this section to

---

[49] The exceptions are cases of causal overdetermination; for example, if two bullets strike a man's heart at once, the man's death has been causally overdetermined.

[50] These cases correspond to what we might call 'normal' cases of causal interruption. For example, to employ the barometer case again, if the barometer were broken and giving a zero reading the closest possible world still would be one in which the air pressure was 30 kilopascals (or something like that) and not a world in which there was no air pressure at all. Contrast this with nomic dependency in which a barometer reading of 0 would imply that there was no air pressure at all. Notice that citing 'relevant conditions' does not support nomic dependency and the strict conditional analysis.

distinguish between these failed objections and the substantive objection which I am raising.

I will proceed as follows. First, I will divide the objections roughly into different categories and outline Lewis's responses to each. Second, I will examine each set of objections more closely and detail Lewis's responses.

The objections which fail divide roughly into the following categories:

1. Objections from the point of view of the assertion that causal dependence is an empirical fact and not a logical consequence of the analysis of counterfactuals.

2. Objections from the point of view that causal dependence is a 'real' property of cause and not simply a function of the way we describe the world.

These first two objections are not unrelated and it is not uncommon to see them both expressed in the same paper. But since they are distinct objections they ought to be distinguished.

3. Objections which question whether correct truth values for particular propositions can be determined by the possible worlds analysis.

4. Objections which question whether any truth values for particular propositions can be determined by the possible worlds analysis.[51]

These latter two objections accept that a logical solution to the problem of causal dependency is possible but question whether the correct solution has been proposed in particular cases.

In this section it will be shown that none of these objections can be employed against Lewis.

The first two sets of objections miss the point completely. Lewis need respond only as follows: if a logical solution solves the problem (as it appears) then there is no reason to argue that the problem has not been solved.

The second two sets of objections are too specific. On a case by case basis the possible worlds analysis works. Given a context in which to evaluate a proposition such an evaluation can succeed. It is only on a global basis that the analysis breaks down. This will be demonstrated in section 6.

Let us now briefly survey examples of these objections and how Lewis can (or does) respond.

1. Counterfactual dependence is an empirical fact and not a logical consequence of the analysis of counterfactuals.

---

[51] A preliminary distinction along these lines was proposed by John Baker.

This set of objections is motivated primarily by epistemological considerations and approaches the subject from two directions:

(i) Counterfactual dependence should depend on some empirical fact, for example, some previously known regularity.[52]

(ii) Counterfactual dependence is assumed because of an asymmetry of knowledge: for example, we 'know' much more about the past than we know about the future.[53]

The answer to (i) has been surveyed above; regularity analyses have been tried and they have been unsuccessful. Lewis remarks, "I think it is time to give up and try something else."[54]

Lewis attacks (ii) as a form of the regularity argument and undermines its premise. The past, he suggests, is as uncertain or undetermined as the future; supposed regularities might not be in fact regularities.[55]

In both cases the response to this set of arguments amounts to the assertion 'empirical evidence does not resolve the problem of causal dependence.' Given that the counterfactual analysis does solve the

---

[52] See Per Thyge, 'On Subjunctive If-Then Statements' for the most straightforward statement of this objection. A variant of this is found in Nute's 'David Lewis and the Analysis of Counterfactuals' where he argues that 'possibility' is not comparable like 'probability'.

[53] Objections of this sort are surveyed in Lewis's 'Counterfactual Dependence and Time's Arrow'.

[54] Lewis, "Causation", p. 557.

[55] Lewis, 'Counterfactual Dependence and Time's Arrow', p. 459.

problem the argument that empirical evidence should be used carries little weight.

Lewis makes this argument clear in "On The Plurality of Worlds". He writes, "Why believe in a plurality of worlds? -Because the hypothesis is servicable, and that is a reason to think that it is true."[56]

2. Counterfactual dependence is a 'real' property of cause and not simply a function of the way we describe the world. Again, we may express this objection several ways:

(i) Causal dependence follows naturally from some previously defined ontology.[57]

(ii) Causal dependence itself is an ontological 'fact' or 'primitive'.[58]

(iii) the possible worlds analysis will determine that causal dependency exists where in fact there is none.[59]

---

[56] Lewis, "On The Plurality of Worlds", p. 3.

[57] See Lewis, 'Counterfactual Dependence and Time's Arrow' for several examples of this. For example, the world might 'really' be indeterminate or there might 'really' be multiple actual futures, etc.

[58] See Jackson, von Wright, (possibly) Davidson, and Goldman for statements of this position.

[59] See Jackson for this argument.

Objections of this sort fail because, once again, they miss the point. Since Lewis's analysis is a logical analysis no ontological assumptions are required.

The distinction between causal and non-causal propositions on Lewis's analysis is a distinction of logical form; should such a form fit (or possibly determine) some particular ontology then all the better but such a fit is not a requirement a priori of the adequacy of the analysis.

Consider Jackson's 'Hume World', for example. According to Jackson, Lewis would have to assert that causal dependency exists in the Hume World even if the Hume world is defined as a world in which there is no causal dependency.

That may be so. But no work could be done - even in the Hume world - by a theory which asserts that there are no causal relations and no causal dependency.

This response is similiar to Lewis's response to the first set of objections.

> It offers an improvement in what Quine calls ideology, paid for in the coin of ontology. It's an offer you can't refuse. The price is right; the benefits is theoretical unity and economy are well worth the entities.... Modal realism is fruitful; that gives us good reason to believe that it is true.[60]

---

60 Lewis, "On The Plurality of Worlds", p. 4. By 'It' Lewis is referring to set theory but his use of the example is obviously meant as a parallel.

3. The possible worlds analysis might not provide the correct truth values for particular causal propositions.

There are numerous examples of this objection.[61] Following Lewis, I will use the following counterexample proposed by Fine:

> The counterfactual "If Nixon had pressed the button there would have been a nuclear holocaust" is true or can be imagined to be so. Now suppose there will never be a nuclear holocaust. Then the counterfactual is, on Lewis's account, very likely false. For given any world in which the antecedent and consequent are both true it will be easy to imagine a closer world in which the antecedent is true and the consequent false. For we need only imagine a change which prevents the holocaust but does not require such a great divergence from reality.[62]

Objections based on this sort of example fall into two categories:

(i) Lewis must posit 'tiny miracles' in possible worlds to both allow them to be as close as possible to the base world and to allow the counterfactual antecedent to be true in the possible world; this breaking of 'natural laws' may skew the selection of the closest possible world and consequently skew the truth values of proposed counterfactuals.

(ii) Lewis must permit 'miracles' which 'repair' the consequences of permitting a counterfactual antecedent to be true in a possible world to bring the possible world to the greatest possible similarity with the actual world; again, this skews the truth values of counterfactuals analysed in this way.

---

[61] Lewis cites Bennett, Bowie, Creary and Hill, Fine, Jackson, Richards, Schlossberger and Slote in 'Counterfactual Dependence and Time's Arrow', p. 467.

[62] Kit Fine, review of Lewis's "Counterfactuals", p. 452.

Against such criticisms Lewis supports the use of tiny miracles to permit the antecedent in a possible world but rejects the use of miracles as a means of repairing such antecedents. Otherwise, as the objection states, the wrong counterfactuals would be true.

The effect of this set of objections is to force Lewis to refine his criteria for the selection of a possible world. Since the proposed counterfactual is true (or supposed to be so) possible worlds in which there is no holocaust given that Nixon pushed the button are not the closest possible worlds to the base world in which Nixon contemplates but does not push the button.

That is - to draw from arguments from previous chapters - the effect of this set of objections is to force Lewis to more carefully define the accessibility relation R.

Lewis's precise criteria are as follows:[63]

[r1] "It is of the first importance to avoid big, widespread diverse violations of law."
[r2] "It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails."
[r3] "It is of the third importance to avoid even small, localized, simple violations of law."
[r4] "It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly."

A fifth rule, which is implicit, is as follows:

---

[63] Lewis expresses these rules explicitly in 'Counterfactual Dependence and Time's Arrow', p. 472.

(v) "There will be no tiny miracle possible which will undo the results of a counterfactual assumption."[64]

The reason is this. For any particular cause there will be typically several effects while for any particular effect there will typically be one cause. Notice how this fits into the logical definition of a causal proposition from the previous section. The counterfactual (-E -> -C) is typically false because there will be several effects (E1, E2, ..., En) of the same cause and even if one of them, E1, is not present in the closest possible world, the others, (E2, ..., En) will still be present and in need of a cause.

These rules are worth stating at length because they are the clearest example in all of Lewis's writings of the dependence of the selection of a possible world on the state of affairs in the base world. We shall return to these rules in the next section.

There is a series of objections in the literature which fit into this category but which are not based on cases such as Fine's examples:

(iii) This (or that) modification to the selection of a closest possible world will give better truth values than Lewis's analysis.[65]

---

[64] Lewis writes, "Divergence from a world such as (G) is easier than perfect convergence to it." Lewis,'Counterfactual Dependence and Time's Arrow', p. 473.

[65] Nute, for example, in 'Counterfactuals and the Similarity of Worlds' suggests examining all worlds which are "similiar enough" to the base world. Bigelow proposes in 'If-Then Meets Possible Worlds' that some possible worlds in which the antecedent remains false should be examined.

Such proposed refinements differ from Lewis's analysis only by proposing that the accessibility relation R be defined in a different way. Some modifications will, perhaps, be accepted; others, such as the Fine objection, will force Lewis to clarify his own definition of R. But none of these objections are objections to the possible worlds analysis of cause, only to Lewis's precise version of it.

4. For any particular proposition, no truth values can be determined.

Again, several formulations are possible.

(i) Possible worlds cannot be studied empirically and thus the truth values of certain propositions which correspond to counterfactual propositions cannot be determined.[66]

(ii) The analysis successfully evaluates only causal propositions which have already been determined to be true causal propositions but cannot be employed to evaluate propositions which are not known to be true causal propositions.[67]

---

66 This view is most clearly stated by Tom Richards in 'The Worlds of David Lewis'. See also Pavel Tichy, 'A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals'.

67 This was my own objection to Lewis during a graduate seminar (1987) and to my knowledge has not reached the literature. The argument was this: suppose some new and apparently causal dependency was observed. It is impossible to know the relevant truth values of corresponding propositions in possible worlds unless the truth value of the proposed causal proposition is known in the base world. That is, more formally, given some proposed O(A->B), it is impossible to know that A->B, B->A, and -A -> -B are true and -B -> -A is false unless it is known that O(A->B) is a true causal proposition.

These objections are closer to the point but too specific. Given a 'context' or base world in which to evaluate proposed causal propositions such an evaluation may succeed.

Objection (i) is answered by Stalnaker's assumption of 'natural systemic coherence' as described in chapter two, section four. Lewis may respond with reference to the rules for the selection of a closest possible world stated above. These rules describe a possible world in terms of their differences from the base world; the possible world will resemble the base world except in cases specifically changed as a consequence of the truth of some counterfactual antecedent in the possible world and not the base world.

Objection (ii) is a bit closer to the point since it attacks the description of the base world itself. Yet it too is too specific. Lewis responds that a true causal proposition is not such if certain propositions are true in some possible worlds; it is such because some certain propositions are true in some possible worlds. Proposed causal propositions for which it cannot be shown that certain propositions in possible worlds are true are not known to be true causal propositions.[68]

Allow me to summarize briefly what has been asserted in this section and show how it leads into the main objection stated in the next section.

---

[68] Lewis gave this response in a graduate seminar, 1987.

I have divided objections to Lewis's analysis of causal dependency into roughly four categories. For each category I have listed a set of objections and outlined Lewis's responses. Not one of the objections in any of the categories succeeds.

The first two sets of objections fail to understand Lewis's program; in effect they amount to the dogmatic assertion that the problem must be solved in one way and not another. Lewis responds by suggesting that other ways have not solved the problem while his own logical solution is effective and suggests that his success where others have failed is a good reason for believing that he is right. Which it is.

The third set of objections amount to no more than proposed refinements to Lewis's accessibility relation R with which possible worlds are selected. These objections carry more force since they do force Lewis to a more precise statement of R. But they do not address whether or not the possible worlds analysis itself is adequate and consequently are much too specific as a general criticism.

The fourth set of objections attacks the process of the evaluation of proposed causal propositions by Lewis's analysis by suggesting that the process yields no results. But the manner in which this set of objections is presented is again too specific; given a context or base world with which to select a closest possible world using R the analysis can indeed select a possible world and allow for the evaluation of proposed causal propositions based on the truth values of corresponding propositions in the selected possible worlds.

There is an ordering, however rough, to the presentation of these objections. We began with objections which completely miss the point and ended with objections which almost make the point but miss slightly. It is a small step from the last objection to the main objection stated in the next section.

To which we now turn.

6. THE FAILURE OF THE POSSIBLE WORLDS EVALUATION OF CAUSAL PROPOSITIONS

The main objection is this:

The selection and description of possible worlds which will be used in the evaluation of causal propositions requires a selection and description of some base world to which the possible world will be similiar. This requirement cannot be fulfilled without first selecting and describing some possible world.

Notice that this objection is exactly the objection proposed in chapter one, section 11, and restated in chapter two, sections 4 and 6.

The consequence of this objection will be that no causal proposition may be evaluated using the possible worlds analysis unless some description of the world is provided prior to any such application of the analysis.

I will proceed with this objection as follows. First I will restate the two problems of evaluation detailed in section 1. Second, I will outline the procedure for solving these problems as explained in section 3 adding, as neccessary, clarifications of the accessibility relation as described in section 4. Third, I will proceed through the steps of the objection as outlined in chapter 1, section 11, showing that either circularity or regress follows necessarily. Fourth, I will show in the context of this objection why the objections listed in section 5 failed.

There were two problems of evaluation.

First. How do we infer that a particular causal proposition is true. As an example we suggested that the evidence which justifies the assertion "Striking a match causes it to light" equally justifies the assertion "Striking a match causes it to have become wet."

Second. How do we infer that a particular proposition is a causal proposition. As an example we suggested that although "If the water is heated bubbles form" and "If the steam rises bubbles form" are equally true counterfactuals only the first corresponds to a causal proposition.

Both these problems are solved on the possible worlds analysis by a precise definition of causal dependency such that a true causal proposition is some proposition in which the effect depends on the cause and the cause does not depend on the effect. Causal dependency

was specified precisely using the possible worlds analysis developed in chapter two.

The definition was as follows.

A proposed causal proposition O(C->E) is a true causal proposition if and only if it corresponds to a set of four counterfactual propositions O(A->B), O(B->A), O(-A -> -B), and O(-B -> -A) such that corresponding instances of these counterfactual propositions may be evaluated such that instances of the first three counterfactual propositions tend to be true while instances of the fourth counterfactual proposition tend to be false.

Accordingly, we will determine whether a proposition is a true causal proposition by examining the truth values of a set of counterfactual statements which correspond to the proposed counterfactual proposition.

Although there will be a number of sets of counterfactual statements of the general form (A->B), (B->A), (-A -> -B) and (-B -> -A) we can focus on just one such set since if the evaluation succeeds with one it will probably succeed for them all. Let us use the general notation stated immediately above to represent that particular set.

Just to keep everything clear let me specify exactly the counterfactuals we will be attempting to evaluate for each problem.

Problem One.

[p1] "If match x is struck it lights"

[p2] "If match x lights it is struck"

[p3] "If match x were not struck it would not light"

[p4] "If match x did not light it would not have been struck"

Problem Two.

[p5] "If water x is heated water x boils"

[p6] "If water x boils water x is heated"

[p7] "If water x were not heated it would not boil"

[p8] "If water x did not boil it would not have been heated"

[p9] "If water x steams it boils"

[pA] "If water x boils it steams"

[pB] "If water x did not steam it would not have boiled"

[pC] "If water x did not boil it would not have steamed"

Recall that a number of instances of each counterfactual will be evaluated (for values of x1,...,xn) under varied circumstances. That is, [p1], [p2] and [p3] will have to be true most of the time while [p4] should be false much more frequently than the preceding. A similiar tendency should be the case for [p5] through [p8] but not for [p9] through [pC].

This determination requires that we obtain truth values for each counterfactual for each value of x. (All possible values? Let's just say a lot.)

According to the possible worlds analysis we will have to select some possible world H for each counterfactual using relation R. The relation R has two arguments, (A,G). The first argument corresponds to the counterfactual antecedent of each counterfactual.

If the antecedent is true in G then R should select G. Presumably both A and B are true in G and therefore the possible world selected for the first pair (A->B) and (B->A) will be G. Since if A and B are true in G then -A and -B cannot be true in G some possible worlds, possibly different, H1 and H2 will have to be selected for each of the latter counterfactuals.

Let us choose two specific counterfactuals which are representative of each: [p1], in which the world selected will be G; [p3], in which the world selected will be some H.

We specify R such that it will choose the correct H. This specification could vary but Lewis has provided us with a precise specification which I detailed while considering the third set of objections in section 5 of this chapter. Let's use that.

Now we are all set. The steps of the objection as outlined in chapter one, section 11, may be followed exactly.

1. We need to select a possible world. That has been established.

2. Closeness is defined by R; we need to find the truth values of some set (defined by R) of atomic formulae or n-tuple relations in

both G and R. Lewis has defined that set in his definition of R (the rules listed in section 5).

3. In some cases these truth values will be truth values of counterfactual propositions in either G or H. This is clearly true if we consider Lewis's definition of the accessibility relation R as stated in section 5. Consider [r1]: we should avoid widespread violations of law. We can imagine what the relevant laws would be for [p3]: for example, laws regarding the reaction of certain chemicals to certain friction. Some such laws are causal laws. Causal laws are evaluated, as noted above, by the evaluation of counterfactuals. Other such laws may be different counterfactuals. Therefore to avoid widespread violations of law we will have to know the truth values of some counterfactuals distinct from [p3] in both G and H. Rules [r2] and [r3] will similarly require that some counterfactuals be known to be true.[69]

4. We need to find the truth values of some counterfactual propositions in either G or H. Q.E.D.

5. To find the truth value of some counterfactual in G go to the first step. We must repeat this entire process for 'chemical x oxidizes when exposed to a surface of a friction coefficient y at velocity v'. To show that this is a causal law (or even a true counterfactual) we will have to repeat steps 1, 2, 3 and 4 which will

---

[69] Cases in which the possible world selected is G, as in the case of [r1], are special cases. These will be discussed in section 7. Note that the objection succeeds whether or not we consider cases like [r1].

result in the need to show that a third causal statement or counterfactual is true. Notice that not just one, but many causal propositions or counterfactuals will have to be true for even one instance of [p3] to be evaluated.

6. To find the truth value of some counterfactual in H we need to find some L which is a possible world defined by R relative to H. If a counterfactual is evaluated in some H then H will become G relative to the evaluation of that counterfactual. L in such a case will be some distinct H'. If H' is the original G then vicious circularity develops. If H is not the original G then infinite regress develops.

Steps 5 and 6 are jump-off points to vicious circularity or regress which will occur for each assessment of each instance of each counterfactual. Even after just one cycle through the loop the number of counterfactuals which will have to be evaluated will be enormous.

As a consequence we see that no evaluation of any proposed causal proposition could ever be completed. Such an evaluation cannot even begin until a possible world is specified and such a specification can never be completed.

Stated in just this way the objection to the possible worlds evaluation of causal propositions becomes clear and obvious.

Why then did earlier objections fail? These objections each addressed only part of the objection as a whole.

Let us begin with the first two sets of objections. A short preliminary discussion is required first.

Recall the list of counterfactuals [p1] to [pC]. Instances of each counterfactual were evaluated for each value of x1,...,xn. These instances varied according to the precise circumstance of the counterfactual. That is, specific details of each counterfactual varied from x1 to x2 to xn. For example, for x1 the match was wet or for x2 the match was struck correctly. By contrast, Lewis's specification of the accessibility relation kept constant certain counterfactuals governing the situation.

Similarity between H and G therefore is on Lewis's account heavily weighted toward a similarity of true counterfactuals. This seems just the opposite to the way it should be given the objection stated above. This intuition - never explicitly stated - is the basis for the first two objections.

Each of the first set and the second set wanted to keep some particular set of specific details constant and vary the counterfactuals. But simply asserting that details should be kept constant does not in itself provide a sufficient objection to the possible worlds analysis. Only the complete objection stated here gives a reason why these details should be kept constant.

The third set of objections is, as I have noted, an attempt to redefine the accessibility relation R. Such an objection might be motivated to define some R such that the selection of some H does not

depend on the truth of some counterfactual in G. But without a clear idea of this motivation such attempts at redefinition have little chance of success.[70]

The fourth set of objections fails to recognize the importance of the third step of the main objection. If it is allowed that some counterfactuals are true in G then rules like [r1] can be employed to define the relation of similarity between G and H. A second step - the recognition that those rules themselves require some sort of possible worlds evaluation - is required to generate the regress and circularity shown in the main objection.

7. FROM POSSIBLE WORLDS TO WORLD VIEWS

In this section I will link the collapse of Lewis's possible worlds evaluation of causal propositions to the development of what I called in chapter one the world view theory.

I shall proceed in two steps. First I shall motivate the need for a number of world views as the only possible response to the main objection as stated in section 6. Second I shall motivate from Lewis's argument a prime thesis of the world view theory in its many forms: that statements are to be evaluated as observation statements only insofar as some theory is employed to provide that evaluation.

---

[70] My own view is that no definition of R will succeed which does not depend on the truth of some counterfactual in G. Certainly, the attempt to define R as a relation of similarity will have to be abandoned. Whether some other R will work remains a question.

The world view theory might be viewed as the only possible response to the main objection.

Notice the need to evaluate more and more counterfactuals which occurs at step 3 of the main objection. Given a termination point at this step (in the form of a set of true counterfactuals which do not need to be evaluated) the regress does not occur.

One way out of such regress will be to stipulate that a number of causal statements or causal propositions are true prior to the application of the possible worlds analysis. But there are no grounds for the selection of such a set. But that is no reason why we cannot simply choose one.

What occurs in the world view theory is that some 'world view' is preselected and then tested against the world. The particular set of counterfactuals is not 'validated' prior to its acceptance but rather 'vindicated' subsequent to its acceptance. In effect the world view theory amounts to the selection of some model H which is subsequently shown to be closest possible to the base world G.

I shall argue that the world view theory suffers from exactly the same problem as the possible worlds theory.

We now turn to the observation-theory distinction.

Consider the counterfactual instance [p1] in the form (A->B) which, along with more problematic counterfactuals, needed to be evaluated

for some causal proposition. The world selected in this instance was the base world G.

Let us assume that we are in G. Then we will have first hand access to the world in which (A->B) is either true or false. Presumably we would determine whether (A->B) is true or false by observation. Let us assume that the world which we observe when we are in G is in fact G.

According to Lewis observation is a form of counterfactual dependency.[71] In particular it may be expressed as a form of causal dependency. Lewis writes,

> This is my proposal: if the scene before the eyes causes matching visual experience as part of a suitable pattern of counterfactual dependence, then the subject sees; if the scene before the eyes causes matching visual experience without a suitable pattern of counterfactual dependence, then the subject does not see.

In the main objection in the preceding objection I argued that either some set of counterfactual propositions will have to be assumed to be true or no evaluation of such a counterfactual will be possible.

World views provide precisely that set of true counterfactual propositions. Accordingly the question of whether someone sees (and consequently what is seen) will depend on that set of counterfactual propositions. Consequently observation is theory-dependent.

---

[71] Lewis, 'Prosthetic Vision and Veridical Hallucination'.

8. AN OBJECTION AND REPLY

The objection is this. During the course of this chapter no distinction has been made between the analysis of causal propositions and the epistemology of causal propositions. It might, for example, be the job of scientists and not philosophers to determine what we know about causation; that is, to fill in the variables in the analysis.

The response is this.

On a meta-level it is reasonable to want to distinguish between analysis and epistemology. On the model-modal theory it is not.

I do not think that on the model-modal theory there is any epistemology outside the analysis nor can there be.

The epistemology occurs at two levels:

[1] The determination of the set K of all possible worlds. Anything that can possible be known about the world is represented in the form of one of these possible worlds.

[2] the selection of some H of K. What is in fact known is a subset of the set of everything that can possibly be known. This subset is represented in one or more possible worlds H or H' which are members of K.

The task of determining what is known is therefore the task of determining what can be known and then selecting some subset of that. The first is clearly an analytical task. What of the second?

Suppose science (loosely described as the determination of what is known) consists of the selection of some possible world.

The selection of some possible world will in turn depend on what is known (since the selection will occur on the basis of some similarity between (salient qualities of) G and some possible world or model H.

In other words, to determine what is known we must first have determined what is known.

That won't work; the circularity is obvious. Notice the exact parallel between this objection and the general objection I am raising against the model-modal theory.

On the model-modal theory I am describing I am suggesting that the thesis is that the selection of some possible world may be established analytically. Now perhaps I am wrong but this certainly is a much stronger thesis than the obviously circular one stated immediately above.

Of course, my objection to the model-modal theory will be exactly the same objection in the end. Even if the attempt is made to select a possible world analytically circularity will still occur. Elaborating this circularity is however much more complex since it requires that the analytical apparatus be dissected.

So I think I can disregard the question of the epistemology - analysis distinction. Suggesting that the selection of a possible world is an epistemological problem causes the theory to self - destruct much more quickly.

And in any case I don't think that most of the model-modal theorists consider such selection to be an epistemological problem.

Chapter Four

REFERENCE AND REPRESENTATION: THE WORLD VIEW THEORY

1. INTRODUCTION

In this chapter I introduce what I have called the world view theory. I highlight important distinctions between the world view theory and positivism and show that these distinctions are the same distinctions as between the possible world theory and positivism.

As I noted in previous chapters, the core idea and the core problem for the model-modal theory is the selection of some model to be employed in the evaluation of some proposition. In this section I show that this remains true for the world view theory.

By 'evaluation' I mean as above; the specification of some value for some proposition. I focus on two types of evaluation: observation-values and truth-values.

By 'observation values' I mean the determination of whether some proposition may be expressed by one or more observation statements. This involves the evaluation of statements themselves and the

determination of whether particular statements are observation statements.

By truth values I mean the determination of whether or not some proposition is true.

As we shall see, on the world view theory such evaluation takes place not by direct reference to the world but rather in some theory or world view. The selection of such theories rather than their methods of evaluation is the most important aspect of the world view theory.

In the possible world theory as described in chapter two there were two primary means of selection: by similarity, and by salient factors determined by context. In this chapter I consider and refute only selection by similarity. The world view theory employs complex theories for selection by context and these will be discussed in chapter five.

This chapter is structured precisely as follows.

In section two I sketch the world view theory in general terms. I show that the focus of current world view theories is in the selection and rejection of different theories and that the evaluation of propositions occurs within the context of such theories.

In section three I show the contrast between the world view theory and positivist theories in terms of the failed observation - theory distinction. In this section I emphasize the contrast between

positivists' attempts to directly evaluate propositions through the use of observation statements and the world view theorists' contrasting indirect evaluation in the context of some theory.

In section four I reinforce this distinction by showing the distinction between referential and representative semantics. I give an informal definition of representation, show that it is employed by the world view theory, and contrast it with reference.

In section five I establish the importance of selection for the world view theory by showing how the distinctions established in sections three and four necessitate the use of 'bridge theories' to select one or another theory. I introduce some complications which result in the use of several such bridge theories.

In section six I show how similarity is employed by the world view theory to select some model. In particular I show how similarity is employed by two current world view theorists, Bas van Fraassen and Philip Kitcher. In addition I show how this means of selection may be described in terms of possible worlds.

In section seven I outline the major objection to selection by similarity. I set up the problem in terms of the evaluation of statements as 'true' or 'observation' statements and then step through the nine steps of the problem as outlined in chapter one, section 11 and employed in several other locations throughout this thesis.

Before moving on to consider the world view theory in detail important to state at this time what will not be issues in this discussion.

First, the issue of which bridge theory, of all those proposed, is best, most appropriate, or most 'true' will not be discussed. The argument will be that all suffer the same defect because of general defects in the model-modal theory itself.

Second, the issue of the ontological status of the world will not be discussed. In particular it will not be discussed whether or not different bridge theories and models imply or correspond to a 'real' or 'actual' world.

As I stated in chapter one and have repeated from time to time through this thesis the ontological status of the world and/or its models is irrelevant to the theory I am describing or the objection I am proposing. Similarily, I have allowed great latitude in the definition of the accessibility relation.

2. THE WORLD VIEW THEORY

In this section I will outline the major features of what I am calling the 'world view' theory.

The general idea is this. Scientific knowledge consists not of accumulating 'facts' about the world but rather consists of one or more global 'theories' or 'world views'.

Accordingly, philosophers of science should not study how certain 'facts' are confirmed or disconfirmed but rather study the forces which lead to the acceptance and rejection of theories.

Contained in this general idea are a number of distinct but related theses. These were outlined in chapter one, section 1. In this section I will state them more precisely.

First. Science should be understood in terms of the acceptance and rejection of 'theories' as a whole and not in terms of the acceptance and rejection of particular propositions.[72]

One of the first proponents of this thesis was Thomas Kuhn who in his "The Structure of Scientific Revolutions" coined the term "paradigm" to represent different beliefs and practices of proponents of different scientific theories.

Kuhn employs the term "scientific revolution" to represent the rejection of one theory and subsequent acceptance of another. Major scientific revolutions include those associated with Copernicus, Newton, Lavoisier, and Einstein.

Kuhn writes,

---

[72] One could be a holist without descending into an anthropological thesis but none of them seem to. Consider Quine (conventionalism), Goodman (linguistic entrenchment), Polanyi (social convention & methodology), Lakatos (progressive problem-shifts), Rescher and van Fraassen (pragmatic virtues), etc. So I'm not sure what it would be like to be a holist without at the same time being an anthropologist. But yes - it is possible - and we probably need such a theory just for contrast.

> More clearly than most other episodes in the history of at least the physical sciences, these display what all scientific revolutions are about. Each of them necessitated the community's rejection of one-time honoured scientific theory in favour with another incompatible with it.[73]

Second. There may exist at any one time not one but several theories each of which may be suitable for a particular application in particular circumstances.

On an over-all basis such theories may be said to be competing with each other but in any case any of a number of theories about the world may be possibly true at any given time. Laudan writes, "The co-existence of rival theories is the rule rather than the exception..."[74]

The idea that different theories may be applicable in different circumstances is traceable to Feigl. Feigl defines four major contexts in which different rules of inference apply.[75]

Modern statements of the world view theory retain this feature. Kitcher, for example, states that different and incompatible mathematical theories may be employed at different times.[76] Van

---

[73] Kuhn, "The Structure of Scientific Revolutions", p. 224. Kuhn meant this only to apply to 'revolutions' and not to 'normal science'.

[74] Laudan, 'A Problem-Solving Approach to Scientific Progress', p. 145.

[75] The deductive, the inductive, the epistemological, and the moral. See Feigl, 'De Principiis Non Disputandum...?'.

[76] Kitcher, p. 158.

Fraassen describes the selection of different theories at different times in terms of theory "acceptance".[77]

Third. Statements are evaluated only from within the context of the theory within with they are asserted.

This has two components. One. Statements are true and false only with respect to the theory in which they are asserted. Two. Statements are 'observation' statements only with respect to the theory in which they are asserted.[78]

Let me address each in turn.

The view that 'truth' is 'truth in a theory' follows directly from the previous two theses listed above. If no one 'theory' is true and

---

[77] Van Fraassen, "The Scientific Image", pp. 4, 8, 9, and many others.

[78] I have of course identified a third component in chapter three - that propositions are only 'causal' propositions within the context of some theory (ie., truth of certain counterfactuals with reference to a possible world.) This is not emphasized on the world view theory although for interesting suggestions in this direction see van Fraassen, "The Scientific Image", chapter 5.

if there is more than one theory then 'truth', if the term is to be employed at all, must be theory-dependent.[79]

This view is perhaps most strongly stated in van Fraassen's "The Scientific Image". Since theories are not themselves true or false but rather 'accepted' on pragmatic grounds, "truth simpliciter does not make sense."[80]

The assertion that a statement is an 'observation statement' only with respect to the theory in which it is asserted is perhaps the most famous and oft-cited thesis of the world view theory. It has its origins in Quine's 'Two Dogmas of Empiricism' but its clearest expression is in Hanson's "Patterns of Discovery".

The idea is this: what we 'observe' is a function of the theory we believe when we make the observation. Hanson writes,

> Let us examine not how observation, facts and data are built up into general systems of physical explanation, but how these systems are built into our observations.[81]

---

[79] This is obviously too quick. The argument is as follows. 1. It is possible that, for some domain D, more than one non-identical theory may be accepted as true. 2. Consider a proposition P, a pair of theories T1 and T2, and a domain D. 3. Let P be true according to T1 and false according to T2 for D. 4. Then P is true of D relative to T1 and false of D relative to T2. 5. Neither T1 nor T2 can uniquely confer truth or falsity to P. (ie., neither T1 nor t2 are true). 6. For any such P there may be a P' such that steps 1 to 5 are true. (This is an inductive step). 7. Therefore truth for any P of any D is relative to some theory T1 or T2. My thanks to John Baker for the essentials of this argument.

[80] Van Fraassen, "The Scientific Image", p. 90.

[81] Hanson, p. 3.

Hanson's classic example is posed by the question, "Do Tycho and Kepler see the same thing in the east at dawn?" In one sense they do, but in another they don't. Tycho sees a mobile sun rising above the horizon; Kepler sees a static sun which appears to change position because of the turning of the Earth.[82]

The collapse of the observation-theory distinction and subsequent rise of 'world view' theories of observation and truth will be detailed in section 3.

Fourth, the selection of which theory to employ in any given circumstance is determined not strictly by its empirical adequacy but also by 'external' factors.

By 'empirical adequacy' I mean a theory's conformity with observational evidence. Exactly what counts as 'observational evidence' varies from theory to theory; hence the need for external criteria.

Kuhn writes,

> Observation and experience can and must drastically restrict the range of admissible scientific belief, else there would be no science. But they alone cannot determine a particular body of such belief. An apparently arbitrary element, compounded of personal and historical accident, is always a formative ingredient of the beliefs espoused by a given scientific community at a given time.[83]

As Kuhn notes, the study of such external factors involves in part

---

[82] Hanson, p. 17.

[83] Kuhn, p. 223.

the study of the history of theory acceptance and rejection. To repeat Suppe from chapter one, section 5,

> What is required is an analysis of theories which concerns itself with the epistemic factors governing the discovery, development, and acceptance or rejection of theories...

Current world view theories reflect this thesis. Lakatos characterizes science in terms of "research programmes which can be evaluated in terms of progressive and degenerating problemshifts."[84] Laudan similarly characterizes scientific progress as successive theories solving more problems than their predecessors.[85] Van Fraassen characterizes theory selection in terms of "pragmatic virtues" related to the "use and usefulness of the theory."[86]

The most important statements of this last feature are theses in which some theory is selected according to some 'context'. This ties directly into the thesis that theories are selected on pragmatic grounds. On this world view theory the presentation of this thesis os complex. I have reserved discussion of this thesis until chapter five.

---

[84] Lakatos, p. 115.

[85] Laudan, p. 145.

[86] Van Fraassen, p. 88.

3. THE COLLAPSE OF THE OBSERVATION-THEORY DISTINCTION AND THE RISE OF THE WORLD VIEW THEORY

In this section I outline the development of one of the key theses of the world view theory, the collapse of the observation - theory distinction and subsequent ascendence of the assertion that what counts as an 'observation' depends on the theory in which it occurs.

The purpose of this section is the following. I want to show that positivists felt that propositions could be evaluated 'directly' by inference from observation statements, and that this attempt failed and was replaced by 'indirect' evaluation of propositions within a theory as a whole. This distinction will play a crucial role in the sections following.

First I shall sort through some distinctions relevant to this discussion. Second I shall discuss verificationism and the important role of distinct 'observation statements' in traditional positivism. Third I shall summarize two key reasons for the collapse of positivist verificationism and outline the emergence of holism. And fourth I shall show that this thesis continues to play a major role in current world view theories.

Here are some distinctions.

Of primary importance is the distinction between an 'observation' and an 'observation statement'. The first is a non-linguistic entity. It has no truth value. The second is a linguistic entity. It has

truth value and can, through a process of inference, confer truth value on other statements.

Related to the primary distinction is a second distinction: that between 'having an observation' and 'describing an observation'. To have an observation is to be in a certain physical state; photons strike the retina of the eye and send nerve impulses to the brain. To describe an observation is to use language to say what some observation is of.

Proponents of the world view theory often argue that both having and describing an observation are required before it can be said that one "observes"; Hanson writes, "There's more to seeing than meets the eyeball."[87] This may be true but it is important to recognize the distinction as 'observations' and 'descriptions of observations' may be used in importantly different ways.

Recognition of this distinction can resolve much of the confusion surrounding interpretations of van Fraassen's theory. On the one hand he says, "To find the limits of what is observable in the world described by theory T we must inquire into T itself..."[88] But on the other hand he says, "I regard what is observable as a theory-independent question."[89] What appears to be a contradiction is not provided that this distinction is maintained.

---

87 Hanson, p. 7.

88 Van Fraassen, "The Scientific Image", p. 57.

89 Ibid.!

Van Fraassen is saying the following. What the limits of observation actually are is not determined by any theory. But any attempt to describe such limits requires that the description occur within the context of some theory.

Having made this distinction I want now to emphasize its importance. In this chapter I am concerned to discuss observation statements and not observations per se. And the claim I am making is that, on the world view theory, whether a statement is an observation statement or not is determined from within the context of a theory.

It is worth noting that a similiar distinction may be made with respect to the 'content' of an observation. Even a person with no theories about anything could recognize that there are different contents to different observations. To describe such contents, however, requires a theory. The nature of such descriptions will, importantly, vary from theory to theory.[90]

Now let us examine the role of observation statements in traditional positivism.

We may describe the positivist theory of verification in terms of three related theses, as follows:

---

[90] See Hans Hahn for one of the first statements of this position. Hanson repeats it in "Patterns of Discovery", pp. 7-15. Van Fraassen is careful as well to maintain this distinction: "It is also important here not to confuse 'observing'... and 'observing that'." P. 15.

[v1] A statement has empirical meaning (and hence can be evaluated) if and only if it is verifiable.

[v2] A person X knows the meaning of a statement S if and only if X knows what difference the truth of S would make to the evidence of X's senses.

[v3] The meaning of a statement S is the difference that the truth of S would make to the evidence of one's senses.[91]

Positivists were forced to distinguish between two forms of verification:

[1] Strong (or direct) verification.

[2] Weak (or indirect) verification.

Statements which are strongly verifiable are or are deduced from observation statements. Statements which are weakly verifiable are not deduced but rather 'confirmed' by appeal to strongly verified statements.

Dancy characterizes weak verification as follows:

Statements (which are weakly verifiable) are equivalent in meaning to a (probably very long) list of statements about what would be observed under different circumstances, all linked by conjunction.[92]

---

[91] These precise formulations are obtained from Dancy, "An Introduction to Contemporary Epistemology", p. 87. Original sources are Schlick, 'Meaning and Verification' and Ayer, "Language, Truth and Logic".

[92] Dancy, p. 89. The first set of parantheses is my own insertion; the second, Dancy's.

Two key problems spelled the ruin of verification.

First, conditional statements about what would be observed in different circumstances cannot all be verified.

Second, it is not clear that conditional statements are strongly verified simply by showing that both the antecedent and the consequent are true.[93]

The failure of verification is therefore directly linked to the failure of positivist analyses of counterfactuals. Just as counterfactuals could not be confirmed directly from evidence in the world, so counterfactual statements about what 'would' be observed cannot be confirmed.[94]

Possible world theorists took one turn at this juncture; they focussed on the analysis of counterfactuals. World view theorists took another turn; they focussed on the need to evaluate theories as a whole.

Quine was the first of the post-positivists to stress this holism.[95] His argument may be summarized as three theses:

---

[93] See Dancy, p. 89.

[94] I am of course using 'verificationism' and 'positivism' interchangably here. No doubt this is a gross oversimplification. The two theories however overlap significantly and - with exceptions noted - I think it is acceptable in this brief presentation to allow such an interchangable usage.

[95] Of course he was not the first ever. Probably the most famous of the holists was Spinoza. Many others followed.

[1] Theory is underdetermined by data. No matter how much evidence we may have there may be more than one theory which adequately explains that evidence.[96]

[2] Non-observational sentences are tested against experience not singly but in groups. Quine credits this thesis to Duhem.[97]

[3] Theory v3, stated above.[98]

What occurs now is that some statement S which is asserted within the context of a theory T may have more or less impact on the empirical content of that theory. To determine the impact of S and consequently its 'observation value' one must consider T as a whole.

A consequence of Quine's theses is the rejection of the thesis of 'reduction'. Quine characterizes reduction as follows: "Every meaningful statement is held to be translatable into a statement (true or false) about immediate experience."[99] According to Quine, such a reduction depended on the ability to sort the 'linguistic component' from the 'factual component' of a sentence. But since the meaning of different terms varies from theory to theory and depends importantly

---

96 See Quine, "Word and Object", chapter 2, and 'On the Reasons for Indeterminacy of Translation'.

97 See Quine, 'Ontological Relativity', p. 90.

98 These theses are stated as such by Dancy, p. 92.

99 Quine, 'Two Dogmas of Empiricism' in "From A Logical Point of View", p. 38.

on extra-linguistic fact this separation is impossible. So, hence, is reduction.

In the modern presentation of the world view theory the empirical content of a theory is a part of and linguistically indistinct from the rest of the theory. The most sophisticated presentation of this thesis is found in van Fraassen's constructive empiricism. Van Fraassen writes,

> We have seen that we cannot interpret science, and isolate its empirical content, by saying our language is divided into two parts. Nor should that conclusion suprise us. The phenomena are saved when they are exhibited as fragments of a larger unity. For that reason it would be very strange if scientific theories described the phenomena, the observable part, in different terms from the rest of the world they describe.[100]

The part of the world which the theory describes which is supposed to be observable is described by the theory itself.

> Science presents a picture of the world which is much richer in content than what the unaided eye discerns. But science itself teaches us also that it is richer than the unaided eye can discern. For science itself delineates, at least to some extent, the observable parts of the world it describes.[101]

Another consequence of Quine's first thesis is that a number of theories may be employed for any given set of observations. This consequence is reflected in van Fraassen and, as noted in section 1 above, other world view theories. Given that observation itself does not determine which of many theories to select other means of selection assume importance.

---

[100] Van Fraassen, p. 56.

[101] Van Fraassen, p. 59.

4. REPRESENTATION: THE WORLD VIEW THEORY AS A MODEL STRUCTURE

In this section I want to show that the world view theory is a model structure as described in chapter one.

I employ two distinct strategies. First, I show that the world view theory has exactly the model structure as described in chapter one. And second I show that the world view theory is distinct from positivism in just the way that the possible worlds theory is distinct from positivism.

To accommodate both strategies I describe the world view theory in terms of 'representations'. By representations I mean entities which are selected on the basis of their relevant similarity to some original and which are employed in order to learn something about that original.

I proceed as follows. First I define very generally what I mean by 'representations'. Then I show exactly what I mean by representations and show that a representation is in fact a model as described in chapter one. Third, I show that world views are representations as defined and conclude that the world view theory has the same structure as the model structure outlined in chapter one.

Fourth I show that the world view theory is distinct from positivism by contrasting representational semantics, which are employed by the world view theory, with referential semantics, which are employed by positivism.

Let me define 'representation' first very generally.

A representation of some object is an entity distinct from that object which is shown to be in some way relevantly similiar to that object. To represent some object is to identify some entity which serves as a representation of that object.

The purpose of representing some object is to learn something about that object by studying the representation of that object.

For example, a photograph of Jill is a representation of Jill. I can represent Jill with that photograph by showing that the photograph is relevantly similiar to Jill.[102]

I can use this representation of Jill - the photograph - to learn something about Jill. For example, I might notice that she is recently tanned (in the photo) and infer that she has gone on a vacation since I have last seen her. Or I may show the photo to someone else who intends to meet her at the airport and needs to know what Jill looks like.

We are all familiar with this concept of 'representation' but if we examine it more closely we can see that it fits exactly into the model structure as described in chapter one.

---

[102] I'll typically argue that there is a causal process from Jill to the camera to the photo - but not necessarily - I might not know the history of the photo but look at it and say - 'That looks like Jill!'.

I shall now outline the crucial points about representation which show this exact fit.[103]

First. Various things can be representations. For example pebbles may represent the solar system or the symbols "cat" may represent a cat.[104]

Second. Whether something is a representation is a function of the actions of human beings. Human beings choose something to represent something else and some entity's status as a representation is a function of that human action.[105]

Third. The thing which is a representation is numerically different from the object it represents. It is often qualitatively different as well, but this is not essential.

Fourth. The representation has various qualities intrinsically in the sense that the presence or absence of these qualities is a

---

103 In this outline I am relying heavily on Jerry Fodor's "Representations". I am, however, generalizing in some important areas since Fodor intends his theory of representation as a theory of mind.

104 It is tempting to say "various things can represent" but that is false - only people can represent.

105 Only as a function of human actions? We might speculate on the limits of what entity may be chosen to represent some other entity. For example, can a dog be used to represent the solar system? But this is a question of the quality of the representation - we prefer that the representation be relevantly similiar to that which it is representing - but if a dog is used to represent the solar system this may be a poor representation but I don't think its poor quality eliminates its status as a representation. So (perhaps tentatively) whether something is a representation is only a status of human actions.

function of the representation and not a function of the object represented.

Fifth. A representation may be more or less 'true' of the original. For example, a drawing of the Eiffel Tower may be a 'good' drawing or a 'poor' drawing. (See my footnote to the second point.)

Sixth. Something may be represented with more than one non-identical representation. Suppose, for example, I wish to represent a cat. I could draw a picture of a cat. Or I could employ the symbols "cat". (In fact I could employ both at once).

Seventh. Relative success as a representation (at least for the purposes of this chapter) is taken as a function of relevant similarity between the original object and the representation of that object. By 'relevant similarity' I mean that a 'better' representation of some original might be some entity which is not identical to the original. For example, suppose we choose to represent the solar system. If we used as a representation an entity which was exactly similiar to the solar system its usefulness as a representation would be limited. We want to use as a representation some entity which approximates the spatial distribution of elements in the solar system but not the size of the solar system.

Eighth. The notion of truth of a representation is to be cashed out in terms of relevant similarity. By that we mean that some statement about some original object which refers to a representation of that object is true if and only if:

(a) the representation is relevantly similiar to the original object, and

(b) the statement is true of the representation.

For example, suppose we look at a picture of Jill and say, "Oh. I see Jill has grey hair." For this statement to be true of Jill, both

(a) the photo is a colour photo of Jill, and

(b) the image in the photo has grey hair,

must be true.

Notice that if the photo is a black and white photo of Jill then the statement is not true - Jill may have red hair. Also, if the image in the photo has brown hair then the statement is also false.[106]

It should be obvious that 'representations', as described here, are exactly parallel to the definition of 'models' as outlined in chapter one.

It should also be obvious that 'world views', as described above, are 'representations'. Here is the argument.

First. Various things can be 'world views'. Usually symbols such as "pv=nrt" are employed, but also drawings or physical models (such as little wooden balls) are employed.

---

[106] Thanks to John Baker for clarifying my earlier attempts at such a description of representation. The structuring of these steps and much of the wording is his.

Second. Whether something is a 'world view' is a function of human actions. No theory or model is intrinsically a world view - it must be accepted by some human to be a world view.[107]

Third. A world view is numerically different from the world. Otherwise it would be the 'world' and not a 'world view'.

Fourth. A world view has certain properties intrinsically and not as a function of the world. Suppose, for example, a world view which is a collection of true and false statements. The truth or falsity of some statement in the world view is a function of its place in the world view and not a function of the world itself.

Fifth. A world view may be more or less 'true' of the world. For example, Newton's world view is generally considered to be less 'true' of the world than Einstein's.

Sixth. There may be more than one non-identical world view.

Seventh. The relative success of a world view is (for the purposes of this chapter) taken to be a function of the similarity between the world view and the world. (In the next chapter we shall see that it may also be a function of the 'usefulness' or pragmatic virtues of the world view).

Eighth. The notion of truth of a world view is cashed out in terms of this relevant similarity. For example, suppose again that some

---

[107] It's impossible to have a 'view' without a 'viewer'.

world view consists of a set of statements. A statement is 'true' of the world if and only if:

(a) the world view is relevantly similiar to the world, and

(b) the statement is true in the world view.

We can conclude the following. Representations, as described in this section, are models, as described in chapter one. World views are representations. Therefore world views are models as described in chapter one.

I have shown how the world view theory is structurally similiar to the possible worlds theory. Now let me reinforce the argument that the world view theory and the possible worlds theory are in fact one theory by showing that they are distinct from positivism in just the same way.

I will do this in two stages. For the remainder of this section I shall show how representative semantics, as employed by the world view theory, differ from referential semantics, as employed by positivism. The second stage will occur in the next chapter where I highlight exactly how positivism differs from the world view theory.

I shall now distinguish representational semantics from positivist semantics.

Positivist theories employed a Tarski - style semantics. The idea was this: truth is defined in terms of the referential concepts of

naming, predication, quantification and satisfaction. For some proposition to be true it had to refer to an object within a specified domain such that the truth of the proposition is satisfied by reference to the domain.

Tarski semantics is therefore a 'direct' semantics. As noted in previous chapters, positivists attempted to explicate the nature of the evaluation of counterfactuals directly, solely with reference to the world. Also noted was the failure of this explication to produce a procedure which would adequately evaluate such counterfactuals.

Successfully to employ Tarski semantics for a theory, the theory must have the resources to enable direct reference to some objects, the objects of which statements in the theory will or will not be true.

As well, and more importantly for our purposes, to save the epistemological programme of the positivists, the theory must have the resources to enable direct reference to observations. That is, the theory must have the resources to express observation statements.[108]

---

[108] This is obviously to gloss over many of the complications of Tarski's theory. In particular, I have not mentioned a primary complication, the distinction between an 'object language' and a 'metalanguage' intended to describe the object language. See Tarski's 'The Semantic Conception of Truth'. For an outline, see Bergmann, et. al., "The Logic Book", pp. 331-342. Van Fraassen discusses and rejects Tarski semantics on p. 67 of "The Scientific Image". Credit is again due to John Baker for this precise formulation.

It was exactly such direct reference to observations which has proven impossible. As I noted in section 3 major arguments were brought against positivism on just this point by some of the originators of the world view theory. Giving up observation statements also led them to give up Tarski semantics and instead to adopt representative semantics.

Representative semantics on the other hand follows more closely Kripke semantics. The truth of a proposition asserted in a world G is determined with reference to a model H selected from a set of models K. This has been described as it applies to counterfactuals in chapter two.

Whether some proposition P is true in some model H does not depend on the truth value of that proposition P in the original world G. Rather, it is determined by the nature of H itself and the selection of some H of K. The relevance of P's truth in H to the corresponding truth of a corresponding P in the original world is determined by the selection of H as the appropriate model to represent G in some particular circumstance.

Notice, then, the exact parallel between the Kripke semantics as described in terms of the model structure outlined in chapter one and representative semantics as described in this chapter.

A second major component of the model structure described in chapter one is the emphasis on the selection of a particular model. It will be in this process of selection that difficulties for the

world view theory begin to emerge. The problem of selection will be discussed immediately below, starting with the next section.

5. SELECTION: AN INTRODUCTION

Most of the work that has been done in the world view theory has focussed on the selection of some theory or world view. There is consequently a wide range of alternatives which will require a judicious sorting out.

The remainder of this chapter and all of the next will be devoted to world view theory approaches to theory selection.

In this section I show the crucial role of the selection of theories in the world view theory. I proceed as follows.

First, I point to distinctions between positivism and the model-modal theory iterated in previous chapters and the sections immediately above. Second, using diagrams I show how 'bridge theories' select particular theories on the model-modal theory but not on positivist theories. Third, I introduce a complication which increased the number of bridge theories employed by the world view theory.

To show the crucial role of selection in the world view theory it is necessary to reiterate the distinction between the model-modal theory (of which the world view theory is a part) and positivist theories.

Throughout this thesis I have shown that positivism asserts, where the model-modal theory does not, two key theses:

[1] Propositions are to be evaluated by direct inference from the world.

[2] Propositions are to be evaluated in isolation without respect to the context in which they are asserted.

These theses have been presented in various forms throughout, expressed one way for counterfactuals, a slightly different way for causal propositions, and differently again for observation statements.[109]

On the model-modal theory contraries to these two theses are proposed:

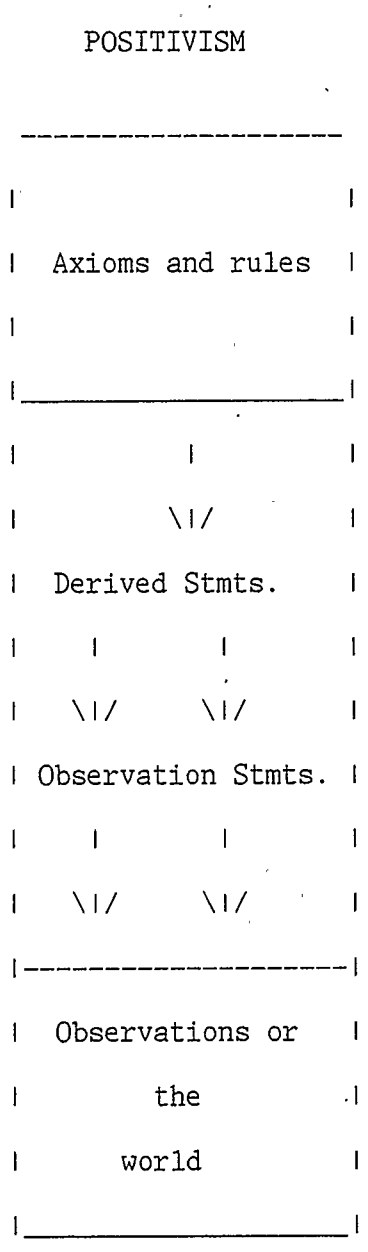
[1] Propositions are to be evaluated indirectly through the use of a model of the world.

[2] propositions are to be evaluated with respect to the context in which they are asserted.

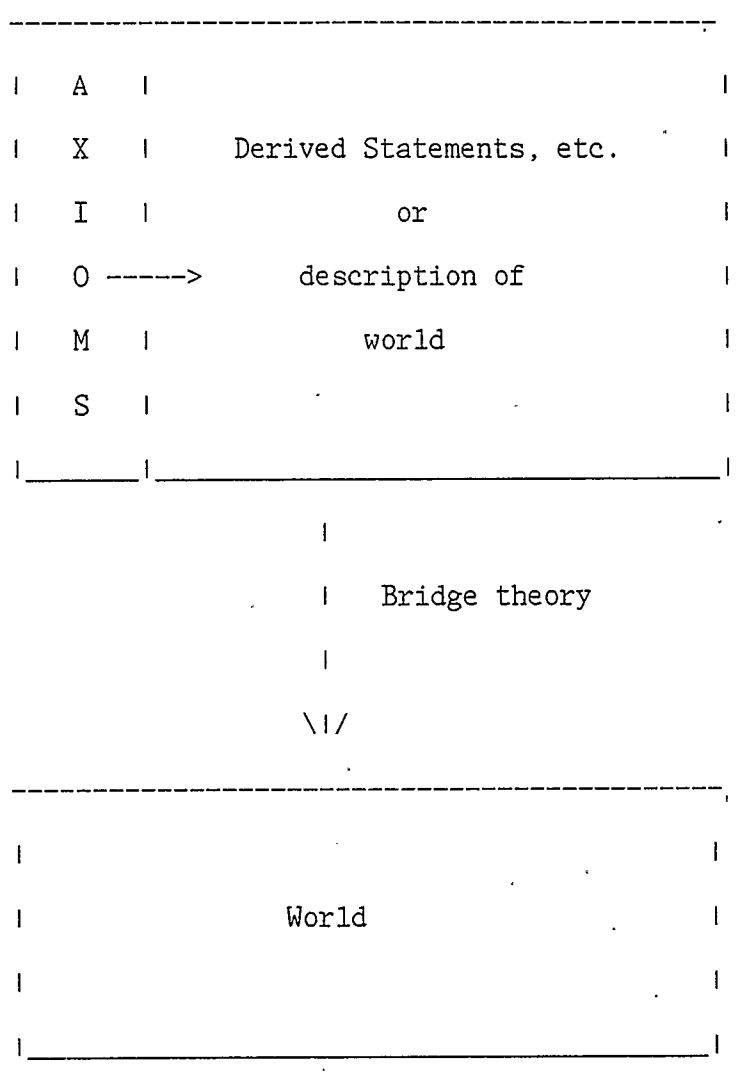
The importance of these two theses may be shown with a diagram which illustrates each of the positivist and the model-modal theory.

---

[109] See chapter two, section 2; chapter three, section 2, and chapter four, section 3 respectively.


POSITIVISM
Axioms and rules
Derived Stmts.
Observation Stmts.
Observations or the world

MODEL-MODAL



In the case of positivism, the theory is directly connected to the world. In the case of the model-modal theory, the theory is not directly connected to the world. (By theory in these diagrams I mean the axioms in conjunction with derived statements, etc.)

In the case of positivism, the semantics of the theory (that is, the analysis of truth conditions and evaluation of truth values) was determined by the direct link from the theory via observation statements to observations or the world.

In the case of the model-modal theory this crucial role is played by what is called a 'bridge theory' which explains in what way statements in the theory are relevant to the world. If more than one theory is proposed, it plays the additional role of selecting which theory will be most relevant to the world at any given instance.

The bridge theory described here just is the accessibility relation R as described in previous chapters.

It has exactly the same role to play: the selection of some model (the theory in question) which is related in some way (as shown by the bridge theory) to the world.

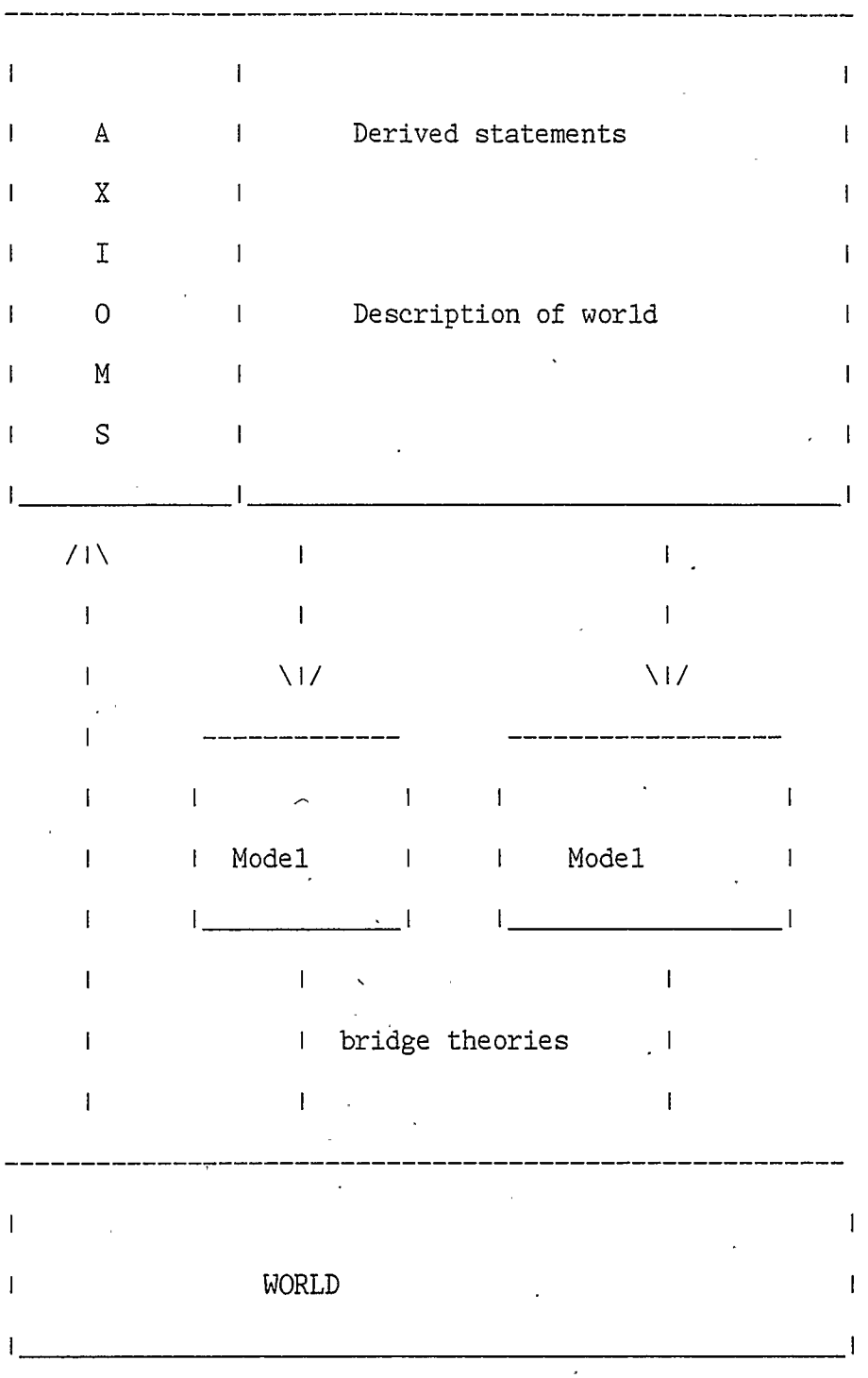
On the positivist theory there is a direct connection between theory and the world and there is only one theory. Correspondingly, the evaluation of statements in that theory does not depend on the context in which the statement is asserted but rather only on the observation statements to which all statements are supposed to correspond.

On the model-modal theory, however, there are no such inherent limitations. Many theories are possible, each of which will be related by the same or different bridge theories to the world in a

different way. Because there are so many possibilities the context in which a statement is asserted is often crucial to its evaluation. Which model is being employed? Which bridge theory is being employed? These questions must be answered before an evaluation can take place.

Allow me now to introduce a complexity which is introduced by more sophisticated world view theories. For any given theory (and there may be many different theories) one or more corresponding models may be produced.

So the complex world view theory more accurately looks like this:



The idea is that the theory is distinct from the models it

generates. But the reader should not be misled. Both the theory and the 'models' are models in the sense described in chapter one. When confusion between these two uses of the word 'model' threatens, as it will, I will employ single quotes (like this: 'model') to show that I mean a 'model' in the world view sense and not a model as I have defined it in chapter one.

An example of this sort of a theory is found in the study of the atom. Bohr's original presentation of the atom was as a theory: the atom was described in words and mathematical equations. Based on that theory 'models' were developed. These 'models' usually consisted of little wooden balls and connecting pegs. The wooden 'models' were connected by bridge theories to two other entities: to the mathematical representation developed by Bohr and to actual atoms in the world Bohr's theory was intended to represent.[110]

It should be noted further that truth values in the 'model' of a theory are independent from truth values in the theory just as truth values in both 'model' and theory are independent from truth values in the world.

---

[110] We should not think that the relation between a theory and its 'models' is direct where the relation between theory or models and the world is not. Bridge theories are necessary to show how the theory relates to the 'model'. For an example of this see Huff, "How To Lie With Statistics" to see how a graph can be incorrectly used as a 'model' of a theory. Van Fraassen uses the Bohr atom as an example, p. 44, and is careful to distinguish a 'model', in which every parameter has a value, to a theory, in which some parameters may have variable values.

We introduce this complication because it shows how several different bridge theories may be employed at any given time. The 'model' of a theory may be selected because it is similiar to some feature of the world. The theory itself may be selected because of some contextual factors.

6. SELECTION BY SIMILARITY

On the world view theory selection by similarity does not frequently occur. In this section I describe how it can sometimes occur.

This section is structured as follows. First I explain why world view theorists often avoid explicit use of selection by similarity. Then I outline a process of idealization which I suggest can lead to selection by similarity. Third I give an example of exactly how selection by similarity is described in terms of possible worlds in one such theory.

Most world view theories do not select theories on the basis of the similarity between the world and the theory. I suggest two reasons for this.

First many world view theorists want to suggest that if a theory is in some way similiar to the world this will be shown indirectly and not directly.

Bas van Fraassen is a good example. An avowed anti-realist, he

does not want to propose that similarity may be established directly. At the same time 'successful' or 'useful' theories will be those which are most similiar to the world. He writes,

> I claim that the success of current scientific theories is no miracle.... For any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive - the ones which in fact latched onto actual regularities in nature.[111]

Second, many world view theorists suggest that the world is similiar not to the theory but rather to some 'model' of the theory.

For example, van Fraassen writes that theories produce 'models' of the world in which there are empirical substructures denoting what is 'observable' on that theory; the theoretical 'observable' is intended to be similiar to what is actually 'observed'.

Van Fraassen writes,

> To present a theory is to present a family of structures, its models; and secondly to specify certain parts of those models (the empirical substructures) as candidates for direct representation of observable phenomena.... a theory is empirically adequate if it has some model such that all appearances are isomorphic to empirical substructures of the model.[112]

The selection of some 'model' on the grounds that it is similiar to the world it describes is often presented by suggesting that the theory is an 'idealization' of the world it describes. Kitcher provides an example of this.

---

[111] Van Fraassen, p. 40.

[112] Van Fraassen, p. 64.

Consider, suggests Kitcher, the Boyle-Charles ideal gas law. Written in full, it is as follows:

(x))(Gx->P(x)*V(x)=R*T(x)

That is, for any x, if x is a gas, then its pressure times its volume will equal its temperature times a constant, R (not to be confused with my use of R to denote the accessibility relation).

Kitcher writes, "If we take the extension of 'G' to comprise the actual samples of actual gases then (the law) is false."[113] But we do not consider the ideal gas law to be 'false'. Rather, as Kitcher suggests, "we can read 'G' as applying to ideal gases, viewing (the law) as true in virtue of the definition of an ideal gas."[114]

That is to say, the ideal gas is a 'model' gas. The ideal gas does not exist on earth; it is strictly a theoretical construction. The question thus becomes (as it has throughout this thesis) why do we select the 'ideal gas' as a model for actual gases?

Notice now that this is the same problem as the selection of a theory as a whole. If the 'theory' postulates some 'model' (as Bohr's theory postulated little wooden balls as a 'model') then we need to show how that 'model' and not some distinct 'model' comes to be selected both for the theory and as relevantly similiar to the world which the theory describes.

---

[113] Kitcher, p. 116.

[114] Ibid.

Kitcher contrasts the ideal gas law with a postulated "horrible gas law", which is stated as follows:

(X)(Hx->[P(x)*T(x)]3/4=R log T

in which the constants stand for the values stated above. Neither the ideal gas nor the horrible gas laws apply to any gases in the world. The question becomes, why select the ideal gas law instead of the horrible gas law.

The selection procedure is based on similarity of the model and actual gases. Kitcher writes, "The stipulation of the characteristics of an ideal gas is warranted by the experience we have of the properties of actual gases. We find that the actual gases approximately satisfy the conditions set down in the Boyle-Charles law."[115]

There are two ways to state exactly what 'approximately' means. First, we can adopt a notion of verisimiltude. Second, we can adopt the possible worlds theory of similarity ordering.

The first way is unsatisfactory. The ideal gas law cannot be 'true' if it only approximates truth. Since we want the law to be 'true' we must reject verisimiltude.[116]

Kitcher adopts the second way. He writes,

---

[115] Kitcher, p. 117.

[116] For a logical statement of this argument see Hempel, 'A Refutation of Verisimiltude'.

> One can naturally think of an idealizing theory as a process in which we abandon the attempt to describe our world exactly in favour of describing a close possible world that lends itself to a much simpler description.[117]

It is tempting to read this as "Every scientific theory is an idealization". This is slightly misleading. It is not a necessary feature of theories that they be idealizations. For example, theories inferred directly from observations are not idealizations. If, however, we reject the possibility of such direct inference then the tendency to make every theory an idealization is that much stronger. This is what happens on the world view theory.

Once on the 'idealization' bandwagon it is very difficult to get off. And once on it is a very short (although often unrecognized) step to a possible worlds theory of similarity ordering.

Given that such idealization occurs, I suggest, a number of theories, each based on a different idealization, are postulated. A method of selection - whatever that method - must be employed and therefore such idealizations fall into the model structure described in chapter one.

A world view theorest might reject the assertion that theories are idealizations selected on the basis of their similarity to the world. That is, the theorist might reject the observation-theory distinction and (perhaps) accept the possible worlds analysis of counterfactuals

---

[117] Kitcher, p. 120.

and yet reject the thesis that theories are selected on the basis of their similarity to the world they are intended to represent.

One way to do this is to employ a pragmatic theory of theory selection. I discuss this in chapter five. But I might note in passing that the precise manner in which selection occurs is not crucial - the objection of circularity and regress may be enforced in any case.

Perhaps the world view theorist might want to reject the thesis of selection altogether (while retaining the other features mentioned above). I do not see how that could be done - we get a series of world views all of equal merit and no means of knowing which to employ in any given case.

The suggestion is this: if there is no direct inference of a theory from the world then there is always a possibility that more than one theory may be suggested or accepted in any given situation. Consequently, either some method of selection is provided or the whole theory is useless.

7. AN OBJECTION TO SELECTION BY SIMILARITY

In this section I show that the objection suggested in chapter one, section 11, may be applied to the selection of some model by similarity as described by the world view theory.

I proceed as follows. First, I set up the problem in terms of observation and truth values. Second, I select which of several possible bridge theories to employ based on the problem. Third, I proceed through each of the nine steps of the objection as outlined in chapter one, section 11.

Let us now set up the problem.

Recall from above that a world view or theory is a model. In this model there may be a further distinction between 'theory' and 'model'. All of these are intended to be relevantly similiar to the world. So when we consider the problem it is important that not one but three bridge theories are required.

We have two types of 'models'. One type of 'model' is what van Fraassen calls an 'empirical substructure' and is intended to determine what statements are 'observation statements'.

The other type of 'model' is similiar to the ideal gas proposed by Kitcher and is intended to determine what statements are 'true' (in this case about gases).

We can treat each of these types of 'model' as the same type of thing: a part of a theory (such that the theory has two parts: 'theory' and 'model').

To simplify matters let's focus on one particular bridge theory. There are three bridge theories to consider.

[1] The bridge theory which selects the 'model' as a model of the 'theory'.

[2] The bridge theory which selects the 'model' as a model of the world.

[3] The bridge theory which selects the theory (and its 'models') as a model of the world.

Selection by similarity is rarely employed for bridge theories [1] and [3]; by contrast it used almost to exclusion by bridge theory [2]. We shall therefore concentrate on [2].

We may now go step by step through the objection as stated in chapter one, section 11.

First. Suppose we want to find the value of some proposition in the world in which the proposition is asserted. To do this we must select some model. On the world view theory as outlined above we need to find some 'empirical substructure' to determine whether the proposition may be expressed using observation statements or we need to find some 'model' such as an ideal gas to find out whether some proposition (such as "gases expand when heated") is true.

Second. Closeness is defined by some relation R, which in this case, is defined as a similarity relation between the model and the world.[118] To find out how close (how similiar) some model is to the

---

[118] Van Fraassen wants exact similarity in terms of isomorphism.

world we need to find the value of some atomic formulae or n-tuple relations in both the model and the world. That is, to show that the 'model' is similiar to the world we must describe both the 'model' and the world.

Third. Some of the atomic formulae or n-tuple relations will be counterfactuals. For example, for observation, at least part of the description will be stated in the form, "If person a were at space-time location x then person a would 'see' structure y." For the ideal gas model statements like "If gas is heated it expands" needs to be true of both model and actual gases.

Fourth. We need to find the value of such counterfactuals in both the model and the world. Otherwise we cannot determine that the values of relevant factors are the same for both model and world and hence we cannot determine that they are similiar.

Fifth. To find the value of such counterfactuals in the world go to the first step. To determine whether people actually see some y when at position x we need some model of seeing. To see whether gases actually expand when heated we need some model of gases.

Sixth. To find the value of such counterfactuals in the model we need to find some further model which is most similiar to the original model. On the world view theory as described above such a model is typically the 'theory' which originated the 'model'. That is, we need to employ bridge theory [1]. Sometimes the properties of the 'model' are determined by the world itself (for example as in cases where the

'model' is made of wooden balls). Thus the world becomes the model for the 'model'.

Seventh. Closeness is defined by R. We need truth values for some propositions in both the original 'model' and the second model (which may be the theory or the world.)

Eighth. To find the truth values of the propositions in the original 'model' go to the sixth step. Vicious circularity occurs as we try to describe the 'model' in terms of itself.

Ninth. The second model is either the world or not the world. If it is the world then go to the first step. Vicious circularity. If it is some other model (such as the theory in question) go to the sixth step substituting the 'theory' for the 'model'. An infinite regress develops.

The world view theory will therefore not be supported by the assertion of similarity between some 'model' of a 'theory' and the world.

Chapter Five

THEORY SELECTION BY CONTEXT: PRAGMATICS AND PRAGMATISM

1. INTRODUCTION

In this chapter I consider world view theories of theory selection determined by the context in which a proposition is asserted.

I show that such methods of theory selection involve the use of salient factors found in the context and show that these factors are described by pragmatic or pragmatist theories acting as 'bridge theories' as defined in chapter four.

Subsequently I outline in detail the various pragmatic and pragmatist theories of theory selection and show that for each of the different theories proposed some description which employs counterfactual conditionals will have to be provided. This sets the stage for the application of the objection outlined in chapter one, section 11.

This chapter is structured precisely as follows.

In section 2 I distinguish between global and local theory selection and show that the same method is employed in both cases.

Following van Fraassen I outline this method in general form and then show how it fits into the model structure described in chapter one. I show how theory selection by context depends on the description of certain salient factors and list some different pragmatic and pragmatist theories intended to determine such factors.

In the next four sections I consider in turn four versions of pragmatism, in each of them suggesting that the version in question essentially involves the use of counterfactuals in its development.

In section 3 I discuss pragmatic theories following from the general definition given by Charles Morris. I distinguish between "pure" and "biotic" pragmatics and then between different levels in each of those. I show for each possible definition of pragmatics that some description in terms of counterfactuals will be required.

In section 4 I discuss pragmatism as defined as a theory about 'practice'. I identify the definition of practice by methodology and by convention. I outline different methodologies and show that any description of a method will require some counterfactuals. I then show that any theory of conventionalism requires some definition of a community; membership in that community must be described counterfactually.

In section 5 I discuss pragmatism as defined as a theory about 'belief'. I consider both how we know that we believe and how we describe the content of our beliefs. For each I suggest that some description using counterfactuals will be required.

In section 6 I discuss pragmatism as defined as a theory about goals. The idea is that the theory selected will be the most useful or successful means to some end. I consider means and ends separately. I show that we must employ counterfactuals to describe both means and ends, and in particular, the relation between the two.

In section 7 I apply the method outlined in section 1 and the conclusions of the next four sections and apply the objection as outlined in chapter one. This application proves successful because every pragmatic or pragmatist theory must provide some description which employs some counterfactuals.

2. THEORY SELECTION BY CONTEXT

On many world view theories the context in which a proposition is asserted will help a person determine which theory is used to evaluate that proposition.

The general idea is that some salient qualities can be determined by context and employed to select a theory. I describe this procedure as follows.

First, I outline two ways in which theories are selected: locally, and globally. Second, I outline a general selection procedure in which salient qualities as determined by context may be employed and show that it fits into the model structure described in chapter one.

And third I outline the options for the selection of such salient features in terms of different theories of pragmatics and pragmatism.

There are two ways in which we use context to determine the selection of a theory.

First, we to use context to choose a theory by specifying a domain of discourse within the world in which a proposition is asserted. Second, we may use context to choose between 'global' theories intended to be applicable over the entire world in which a proposition is asserted.

Let me give examples of each.

In the first case some 'domain of discourse' is specified which may have background theories distinct from theories covering other such domains.

For example, Hanson writes,

> Consider how the cause of death might have been set out by a physician as 'multiple haemorhage'(sic), by the barrister as 'negligence on the part of the driver', by a carriage builder as 'a defect in the brakelock construction', by a civic planner as 'the presence of tall shrubbery at that turning'.[119]

For each person the context in which the death was viewed was different; each person viewed the death in the context of some particular occupation and evaluated the cause of death in accordance with a background theory suitable to that occupation.

---

[119] Hanson, "Patterns of Discovery", p. 54. Cited by van Fraassen, "The Scientific Image", p. 125.

This sort of local evaluations of cause may or may not contradict each other. The point here is that each is selected for a particular case from a particular point of view.

Examples of the second sort are of the replacement or competition between global theories.

Kuhn's theory of paradigm shifts is the classic case of this sort: Copernicus's suggestion that the earth moves dramatically revised the predominant Ptolemaic 'world view' of the time.

Polanyi characterizes the selection of such global theories in contextual factors such as personal preference, community standards and accepted scientific methodology. For example, Polanyi writes that, "Copernicus gave preference to man's delight in abstract theory..."[120]

Similarily, the successes and failures of competing theories are described not by the theories' empirical successes alone but rather in the context of social conventions of evidence, assessments, or success at solving (socially defined) 'problems'.[121]

Whether theory selection occurs locally or globally the idea is the same: the theory selection is governed by some context in which a proposition is asserted.

---

[120] Polanyi, "Personal Knowledge", p. 3. Polanyi characterizes community standards and methodology in terms of 'skills' which are implicitly passed on to students.

[121] See Laudan and Lakatos.

We now examine how this selection occurs.

In brief, the answer is as suggested in chapter two: theory selection is determined by a list of salient factors; these factors are determined by the context in which a proposition is asserted.

Salient factors are those relevant to the evaluation of some proposition or set of propositions. In chapter two examples of salient factors were "Caesar's ruthlessness" and "Caesar's primitive technology". Immediately above the salient factors mentioned were the occupations of the persons involved, social convention, methodology, and personal preference.

Lewis does not specify exactly how such salient factors are selected by context. Van Fraassen suggests a general procedure in "The Scientific Image". We shall now examine that methodology for, as we shall see, it fits precisely into the model structure described in chapter one.

According to van Fraassen the context in which a proposition is asserted determines a 'contrast class' and a 'relevance relation'. The idea is that these two factors select a 'background theory' within which a proposition may be evaluated.[122]

The parallel with the model structure of chapter one is strikingly obvious. The 'contrast class' is a range of "alternative

---

[122] See van Fraassen, p. 129: "The context... determines relevance.... This contrast-class is also determined by context." The "background theory" is related to context on page 145.

hypotheses"[123] analogous to the set of models K as defined in chapter one. The relevance relation (which performs the selection) is analogous to the relation R defined in chapter one. To complete the picture van Fraassen specifies a "topic", Pk, which is analogous to (a part of) the base world (original) G.

This parallel is reinforced by van Fraassen's own formulation. Employing the symbols Pk for the topic, X for the contrast class and R for the relevance relation a question Q is described by the triple:

Q=<Pk,X,R>

which is exactly analogous to the triple described in chapter one.[124]

There are two aspects to context which determine this selection. First, within the proposition P there may be an implicit presupposition of some background theory or facts. Second, the context in which P is asserted may include some presupposition of background theory or facts.

Van Fraassen suggests some obvious examples of this. For example, if someone asks the question, "Why is the conductor warped" then implicit in this question is the statement, "The conductor is warped". If the question is asked at a power plant then a background theory of electricity, magnetic fields and force would be assumed.

---

[123] Van Fraassen, p. 129.

[124] Van Fraassen, p. 141.

Factors such as this are "variables" determined by context. The close connection between the model structure, the possible worlds theory and the world view theory becomes even clearer. Van Fraassen writes,

> Among such variables will be assumptions taken for granted, theories accepted, world-pictures or paradigms adhered to, in that context. A simple example would be the range of conceivable worlds admitted as possible by the speaker; this variable plays a role in determining the truth-value of modal statements in that context, relative to the 'pragmatic presuppositions'.[125]

To reiterate, the idea is this. Various "variables" determined by the context act to determine which of a "range of hypotheses" will be selected. This fits exactly into the model structure described in chapter one.[126]

---

[125] Van Fraassen, p. 137.

[126] This simplifies van Fraassen's theory slightly. There are two further factors worth mentioning. First, van Fraassen characterizes the set of alternative hypotheses as propositions; for example, the alternative hypothesis to "the conductor is warped" is "the conductor is not warped". Other features of the 'worlds' in which these alternative hypotheses are true are compared to the 'world' in which the original proposition is true and some feature present in the latter world (van Fraassen calls it K) A provides an 'answer'. This explains why van Fraassen represents the 'base world' with a proposition Pk and not some symbol which denotes a 'world' such as G. The model structure I have employed in chapter one does not specify that members of the set K must be worlds; therefore this complication does not pose a threat. Second, van Fraassen characterizes 'relevance' in terms of increased probability; that is, $P(Pk)<P(Pk/A)$. The proposition is more likely to be true if some answer A is true because in the 'world' in which the answer A is true so is Pk, and A is not true in other members of X (the set of alternatives) and Pk is not true in those (by definition). This is a refinement on the Lewis-Stalnaker version. Recall that in G a proposition P is true if P is true in some possible world H. Here a proposition P is more probable if P in G if P is true in H. So again, no problem is posed.

To assess this description of how context determines theory selection it is necessary to specify the "variables" involved and consider whether they will provide the necessary selection. I will suggest that for all such postulated variables the problems of circularity and regress occur.

The range of such variables is covered by various theories of pragmatics and pragmatism. To conclude this section I will outline the different options. In sections subsequent I shall examine each in turn.

As noted there are two major theories: theories of pragmatics and pragmatism. The two are not unrelated but they are distinct. On the former, the impact of different presuppositions on the language of discourse is assessed. On the latter, pragmatism, various criteria for theory selection are specified.

There are two major schools of pragmatics, described by Morris as the "pure" and "biotic" aspects.[127] Pure pragmatics defines presupposition and implicature inherent in the language itself. Biotic pragmatics studies various features of the language user (such as psychological state or socialization).

There are three major schools of pragmatism, each characterized respectively by the philosophies of Peirce, James and Dewey. First, pragmatism is defined in terms of 'practices'. Second, pragmatism is

---

[127] Morris, "Foundations...", p. 30.

defined in terms of 'beliefs', and third, pragmatism is defined in terms of 'goals' or 'ends'.[128]

In the sections to follow we shall examine each of the definitions of pragmatics and pragmatics with an eye to theory selection and the main problem of chapter one, section 11.

In each case the method I shall employ will be the same. I shall outline the theory and some of its major variations and show how it is employed in the world view theory of theory selection. I shall then identify in each some factor which must be described and, more importantly, described counterfactually.

It is with this last that I shall be able to, in a concluding section, employ the main argument of chapter one, section 11.

3. PRAGMATICS

Pragmatics is concerned with the impact of theoretical presuppositions on language. As noted above, there are two distinct schools: the school of "pure" pragmatics; and the school which studies the "biotic" aspects.

In this section I shall examine each of these schools in turn and show that in both some description in counterfactual form is required.

---

[128] These are obviously oversimplifications. As early as 1908 Lovejoy was able to distinguish thirteen forms of pragmatism, although to be fair about it, he distinguished pragmatic theories of meaning, truth, knowledge and ontology.

For the biotic school I shall first outline the range of topics it covers and then, following Martin, define three 'levels', each of which will be shown to entail counterfactuals of some form.

For the "pure" school I shall distinguish between two types and argue that each consists of a description of some language in some other metalanguage. Rules of translation of or reference to the object language will be required and, I argue, some of these rules will be counterfactuals.

Finally, I examine van Fraassen's apparently non-counterfactual description of pragmatics and show how even a simple example requires some description in terms of counterfactuals.

We shall now consider the "biotic" school of pragmatics. It is characterized first by Morris. He writes,

> It is a sufficient characterization of pragmatics to say that it deals with the biotic aspects of semiosis, that is, with all the psychological, biological, and sociological phenomena which occur in the functioning of signs.[129]

This list may be extended almost indefinitely. R.M. Martin writes,

> In pragmatics we take account not only of the syntactical and semantical features of language but also one or more of the following: the users of the language taken individually or severally or as members of social groups, the mental states or brain states of the users as well as their activity or behaviour as correlated with their use of language, the physical, biological or social circumstances in which expressions of the language are used, the purposes for which they are used, etc.[130]

---

[129] Morris, "Foundations...", p. 108.

[130] R.M. Martin, "Toward A Systematic Pragmatics", p. 9.

Martin characterizes three levels of pragmatics which will be useful: first, relations between the language and the user such as 'acceptance' or 'belief'; second, relations between users' actions and responses and linguistic stimuli; and third, various social features of language.[131]

Of these the level of 'acceptance' or 'belief' is of most interest to the world view theorist. We may, however, show that problems for the world view theory may be generated at all three levels.

First, how do we determine that the following statement is true?

[p1] X accepts a at time t.

As Martin suggests, "The experimenter asks X at time t whether a holds."[132] But preconditions must be set: a certain "lapse of time" must be allowed for a response; the circumstances of the experiment must be "normal"; evidence from other data, such as actions, must not contradict X's response.

Accordingly, even at the most primitive level of 'acceptance' a counterfactual stating the prior conditions for such a test must be specified. Call the preconditions 'C' and call X's response 'R' and X's 'acceptance' of a at t 'A'. Then the counterfactual is:

[p2] (C&R)->A

---

131 Ibid.

132 Martin, p. 10.

At the second level of behaviour and responses to linguistic stimuli a similiar counterfactual may be established. To state that X responds with some behaviour B in response to stimuli S relevant preconditions must be set such that X does not have any prior mental attitudes which would cause such behaviour in the absence of S such that, if S had not occurred, B would not have occurred either.

Accordingly we get the pair of counterfactuals:

[p3] (S&C)->B and -S -> -B

Notice the parallel between this pair and the pair discussed in chapter three for causal propositions. The reason is the same: we want to specify that some linguistic stimulus causes some behaviour.

At the third level a description of social conditions is required. Some of this description will not involve counterfactuals but much of it will involve relevant counterfactuals such as:

[p4] If X utters the F word in situation s then the police will arrest X.

Therefore in all three levels of biotic pragmatics reference to some counterfactual description is entailed.

We now examine "pure" pragmatics.

We may distinguish between two types of "pure" pragmatics: what I shall call "pure pure" pragmatics, and "natural language

pragmatics".[133] By "pure pure" pragmatics I mean the study of formalized language systems such as logic, or mathematics. By "natural language pragmatics" I mean the study of natural languages such as English or German.

In both cases a description occurs; the relevant distinction is only in what is being described. In both cases a similiar method is adopted: a metalanguage is defined such that all the symbols of the 'object' language are (the equivalent) of nouns in the metalanguage.[134]

Features such as presupposition and implicature are then described in the metalanguage.

To create a metalanguage from some object language a set of 'translation rules' is required. Many of those are non-conditional (such as rules of identity) but many of the important rules are expressed in conditional form - rules involving designation and truth conditions. For example,

[p5] If X is a formula of M, it is a formula of SM(t)(des).[135]

---

[133] This distinction is originally Carnap's; in his terms the two types are "pure" and "descriptive" pragmatics. See Carnap, 'Meaning and Synonymy in Natural Languages'. Cited in Martin, p. 3.

[134] For example, R.M. Martin characterizes object languages following closely Tarski's formalization of a theory of types and from there defines a metalanguage similiar to Tarski's M.

[135] This example was taken somewhat at random from Martin, p. 30. There are many other examples.

Therefore pure pragmatics also requires description in terms of counterfactuals.

Let us now reconsider van Fraassen's use of pragmatics.

Recall from above that if someone asserts the question "why is the conductor bent" the presupposition was that "the conductor is bent". The determination of this presupposition appears to be non-conditional unless it is closely examined.

According to van Fraassen, the presupposition can be deduced only if there is a direct answer to the question. If there is no direct answer (suppose the conductor was not in fact bent) then there is no presupposition.[136]

So we get the counterfactual:

[p6] If there is a question and it has a direct answer then there is a presupposition.

We could stop here but let us take this one more step.

[p7] If there is a question (why is the conductor bent?) and it has a direct answer (ie., if it is bent) then there is a presupposition (that it is bent).

Which gives us the absurd:

---

[136] Van Fraassen, p. 138.

[p8] If there is a question and the conductor is bent then there is a presupposition that the conductor is bent.

There is therefore no direct presupposition. Presupposition depends on there being a direct answer which depends on the 'facts of the matter'. Such facts will invariably require a description and this description will often be in terms of counterfactuals.[137]

4. PRAGMATISM AS PRACTICE

C.S. Peirce defines truth in terms of belief and belief in terms of 'practice' or 'actions'.

"Belief is only a stadium of mental action, an effect on our nature due to thought, which will influence future thinking."[138]

In this section I shall consider different ways of defining pragmatism in terms of practices. I proceed as follows.

First, I identify two major ways to characterize practices: by methodology, or by convention. Second, I examine methodology, show three ways in which it may be defined (using examples of each) and

---

[137] This argument is obviously too quick but I am inclined to believe that a closer inspection would show more counterfactuals, not fewer. It should be of course noted that there will be same cases in which presupposition can be established directly, such as in cases like "Suppose A is true, then..." in which the presupposition is explicit.

[138] Peirce, 'How To Make Our Ideas Clear', p. 121.

show further that each of these three ways involves the use of counterfactuals. Third, I examine the definition of practice by convention, argue that convention requires a community, and show that any description of a community will require the use of counterfactuals.

According to Peirce, human thought and ideas (should) depend on a method of signs, developed through practice, which provides socially standardized ways in which we refer and describe. Truth, for Peirce, occurs when a 'fixity of belief' occurs such that the social standards are themselves fixed.[139]

There are therefore two ways in which we might interpret Peirce's pragmatism: as a theory of methodology, and as a theory of community conventions. Although distinct the two are not unrelated. However we may deal with each separately.

Let me first consider pragmatism as 'methodology'.

Theories of methodology rarely occur in isolation. Polanyi suggests that methodology is determined by social convention. Ackermann suggests that methodology is defined in terms of instrumentation (what instruments we use). Rescher suggests that

---

[139] Peirce's 'truth in the long run' has been subjected to a steady stream of criticism by world view theorists. See Rescher, "Peirce's Philosophy of Science", pp. 4-6 for a summary.

methodology may only be understood in terms of some 'purpose or goal'.[140]

I will consider the question of the conjunction of one form of pragmatism with another below. So to focus the discussion, let us consider practice in isolation.

There are different kinds of practice. For example, 'practice' as intended by Peirce meant the use of signs. Practise, as defined by Polanyi, is defined in terms of 'skills'. Practise as defined by Rescher is in terms of our "employment" of a theory.[141] These three definitions form a useful spectrum and can be considered representative.

Let us consider each of these definitions in turn.

If practice is defined as 'the use of signs' then it resembles closely pragmatics as described in the previous section. The primary distinction is this. Pragmatics determines linguistic practices subsequent to the use of a language. Peirce's definition of practice stipulates linguistic practices prior to the use of a language.[142]

---

[140] Polanyi, "Personal Knowledge", p. 53 (Tradition) and elsewhere; Ackermann, "Data, Instrument and Theory", Rescher, "The Primacy of Practise", p. 3.

[141] Polanyi, "Personal Knowledge", p. 49.; Rescher, "The Primacy of Practise", p. 1.

[142] This distinction is not hard and fast, of course. Some aspects of "pure pure" pragmatics require a stipulative language while some aspects of Peirce's definition require existing linguistic conventions.

If defined prior to the use of a language then linguistic practice will be defined as a set of rules. A good example of this are the rules of symbolic logic. This allows us to establish the existence of counterfactual description on two levels.

First, some of the rules themselves will be counterfactuals. Modus Ponens, for example, reads "If A, and if (A->B), then B".

Second, rules which specify the conditions under which the rules are to be employed will be counterfactuals. For example, some rule might read, "If confronted with a new phenomena apply the rule of designation to name that phenomena."

Therefore if practice is defined as the prior stipulation of a use of signs a description of that stipulation will be required in the form of counterfactuals.

The second definition of practice is in terms of skills. Among those skills will be the use of signs, as described above. Other skills will include the manipulation of tools or instruments.

On this definition there might not be an explicit specification of the rules. Polanyi writes, "I shall take it as my clue for this investigation the well known fact that the aim of a skillful performance is achieved by the observance of a set of rules which are not known as such to the person following them."[143]

---

[143] Polanyi, p. 49. Ackermann, in "Data, Instruments and Theory" expresses a similiar proposition for the use of scientific measuring devices (such as rulers).

In the question of theory selection, however, reference to a set of rules which is not known will be difficult. The rules of theory selection are described only when such implicit rules are made explicit. This is in fact exactly Polanyi's method.

Accordingly the same counterfactuals as generated for the use of rules governing signs can be generated for the use of rules governing skills: rules themselves which are counterfactual, and rules for the use of rules which are counterfactual.

On the third definition practice is defined in terms of our 'employment' of a theory.

Once again the rules for such employment will in some cases be counterfactuals, and the rules for the use of such rules will be counterfactual.

I will not give examples of such rules. The reason is this. The rules for the employment of a theory are just those rules for which we are searching. It would be somewhat circular to define the rules for the employment of a theory as: rules for the employment of a theory.

This third definition thus has obvious difficulties.

The second major aspect of Peirce's pragmatism is the stipulation of some practice by convention. Let us now discuss this aspect.

There are two major questions concerning 'convention'. They are as follows:

First, who's convention is it?, and

Second, what is the convention?

We have discussed the second question immediately above. Let us consider only the first.

On Peirce's theory the conventions which counted were conventions held by the "ongoing community of inquirers". This community, according to Peirce, is defined by its members being "intellectual".[144]

Members of the community will therefore have to be described in some way such that, if they have some particular quality, then they are "intellectual". The list of qualities is potentially enormous but a listing will not be required. The relevant counterfactuals have been generated.

[p8] If person X has quality Y then X is "intellectual".

And:

[p9] If X is "intellectual" then X is a member of the "community of inquirers".

Specifications of a "community" need not be restricted to persons of some or another specific quality but might be quite general: the

---

[144] Peirce adds that extraterrestial beings should be considered members of this 'community' provided they are intellectual.

community of language users (for linguistic entrenchment), the community of epistemic beings (for truth by convention), etc.

But we want some specification of the community such that, for example, it includes persons and excludes rocks and trees. And any such specification will have to be of the form "If X is A then X is a member of B".[145]

Allow me to summarize this section. We distinguished between two aspects of "pragmatism as practice": practice or methodology, and practice as determined by convention. In the former we identified three types and showed that for each type a specification of the particular type involved a description in terms of counterfactuals. In the latter we asked the question, who's convention?, and decided a community is needed, a community which is definable only by means of counterfactuals.

We must conclude therefore that any definition of "pragmatics" as "practice" will involve a description using counterfactuals. Accordingly, all definitions of pragmatism as 'practice', when employed in the selection of some theory, will fall to the main argument of chapter one, section 11, as I shall demonstrate below.

---

[145] We might in fact want to become more complex and explicitly rule out some entities or persons: if X is not an a then X is not a member of B. Our criteria might be quite complex and require a series of counterfactuals. It is worth mentioning that we want to employ the contrary-to-fact form of counterfactuals; that is, we want the community to include potential future members as well.

5. PRAGMATISM AS BELIEF

In this section I consider pragmatism as defined as 'belief' in a theory. For the purpose of selecting a theory the stipulation might be that the theory selected is a theory which is believed.

As in the previous two sections my major intent is to show that various ways of interpreting this definition of pragmatism all require some sort of description in the form of counterfactuals.

I proceed as follows. First I outline James's theory and show the motivation for defining pragmitism in terms of belief. Second I distinguish between two major aspects of this definition: first, how we know that we believe; and second, how we specify the content of our beliefs. Third, I show that we know that we believe in terms of the satisfaction of some set of conditions, thus showing the need for a counterfactual. And fourth, I show that the content of our beliefs must be described using counterfactuals.

William James, like Peirce, characterized truth in terms of practice. But unlike Peirce, who defined pragmatism in terms of a general methodology, James defines pragmatism as actions in response to specific options.

Some such options are, according to James, both lively (we have a choice) and forced (we must make a choice). Often there is no

validation of such a choice; we act on faith. Pragmatism is thus defined in terms of this faith or belief.[146]

James writes, "'The true', to put it very briefly, is only the expedient in the way of our thinking, just as 'the right' is only the expedient in our way of behaving."[147]

James is concerned primarily with two aspects of belief: first, how is it that we know that we believe some particular fact; and second, what is the content of our beliefs.

Let me deal with each of these questions in turn for in answering these questions I will generate the requisite counterfactuals.

According to James we know what we believe because we have a 'sentiment of rationality'.[148] We sense an 'emotional preference' which causes us to choose between options of equal standing. It is this emotional preference which thus governs belief.

James's emotional preference can be applied to the selection of theories as well. For example, we have a preference for theories

---

[146] That is to say, our beliefs are underdetermined by the evidence.

[147] James, "Essays in Pragmatism", p. 170. It should be noted that James wants beliefs to be chosen with respect to our 'ends' or 'goals'; since I discuss ends and goals in the next section I will not discuss it here.

[148] James, "Principles of Psychology", Volume One, p. 4.

which predict the future.[149] As well, notes James, such theories have an obvious utility: they help us survive.[150]

The problem is therefore one of characterizing 'emotional preference' and relating such preference to the selection of some belief in some theory. Possibly this characterization will be in the form of counterfactuals.[151]

But even if not, the requisite counterfactual has been generated:

[p10] If person X has emotional preference p then X will select theory T.

Perhaps we could characterize 'belief' more directly. We could either characterize belief in terms of some physical state or in terms of some disposition (as Ryle does). The latter is obviously counterfactual. The former is counterfactual as well:

[p11] If person X is in physical state p then X will believe (select) theory T.

---

149 James, "Principles of Psychology", p. 14.

150 See the next section.

151 For example, see Ryle, "The Concept of Mind", in which such 'mental states' are described as dispositions to act. This seems in accord with James's theory as well: the only useful cognitions are those which lead to action. See "The Principles of Psychology", pp. 18-20.

Notice that [p11] is just a generalized form of [p10]. Any characterization of belief in terms of some condition result in the use of some form of counterfactual.[152]

Now let us consider the content of belief. This may seem straightforward: what is believed is the theory selected. But it is not.

Van Fraassen, for example, distinguishes between a belief that a theory is 'true' and a belief that a theory is 'empirically adequate'.[153] Accordingly we may distinguish between: first, belief in the content of a theory; and second, belief in the status of a theory. We may further distinguish within the first option: belief that the content of a theory is true, belief that it is empirically adequate, etc.

We can summarize these distinctions as follows: belief in some entity (be it a theory, a content of a theory, a proposition, etc.) is not belief simpliciter; rather, it is belief that such entities have some status.

The counterfactual follows very quickly. If we believe that some entity has some status we must specify what the conditions are for it

---

[152] I am not inferring from one example to a generalization here. Rather, the example demonstrates an instance of the generalization. The generalization follows from the argument that there must be some conditions under which it may be said that X believes Y.

[153] Van Fraassen, p. 12.

to have that status. We then assign the status on the basis of the occurrence or non-occurrence of those conditions. Accordingly we get counterfactuals of the form:

[p12] If entity P satisfies conditions C then P has status S.

Or the converse:

[p13] If P does not satisfy C then P does not have status S.

Probably both.[154]

In summary, the conclusion of this section is exactly the same as the two preceding. No matter how pragamtism is characterized in terms of belief, if it is so characterized, some description will be required which includes some counterfactuals.

6. PRAGMATISM AS USEFULNESS

This final characterization of pragmatism is also probably the most popular. A theory is selected because it is useful, it works, and this usefulness is defined in terms of some end goal or desire.

---

154 Again we can make this more complex. If, for example, there are many truth conditions which a proposition must satisfy in order to be true then the proposition will have to satisfy a series of counterfactuals in order to be true. Note that this part of this section is closely related to the epistemological question of 'grounds for belief'; there is a large body of literature which suggests that certain beliefs must satisfy certain conditions in order to be 'true', etc.

Exactly how such a characterization is to proceed has been the subject of a wide-ranging philosophical discussion. In this section I shall consider only the major alternatives; variations will suggest themselves.

I proceed as follows. First I outline Dewey's theory and distinguish between the usefulness of some means and the ends toward which that means is applied. Second, I discuss ends and show that ends must be described and the conditions for the satisfaction of such ends must be described. I then apply this argument to a specific discussion of specific ends which I distinguish as 'practical' and 'cognitive' ends.

Third, I discuss means. I distinguish between describing means and describing means as 'means to an end'. The former I discussed in section 4. For the latter I show that some process of vindication is required to show that a means is a means to an end; such vindication proceeds inductively and therefore is described counterfactually.

I conclude that if theory selection is to be explicated in terms of usefulness toward some end then some counterfactual description will be required.

Let me first outline (very briefly) Dewey's theory.[155]

Dewey describes pragmatism in terms of scientific inquiry. For Dewey, inquiry occurs as a consequence of doubt. Such doubt is for Dewey inherent in the nature of existence and the world.

> It is the situation which has these traits. We are doubtful because the situation is inherently doubtful.... The indeterminate situation comes into existence from existential causes, just as does, say, the organic imbalance of hunger.[156]

To respond to these problems - and we must, if the problem is something like hunger - we inquire into the nature of the indeterminacy. Dewey proposes logic as the method of inquiry. The idea is to state the problem precisely. From this precision possible solutions may emerge.

From this brief characterization it is possible to see different aspects of the definition of pragmatism as usefulness. I shall now do so.

The clearest and most obvious distinction is that between means and ends. We see in Dewey this distinction: the ends, which is the solution to some problem; and the means, the use of logic as a theory of inquiry.

---

155 I will limit my discussion to Dewey, not because later work in this tradition has not been useful and interesting, but because the different approaches to pragmatism as usefulness are most clearly seen in Dewey, as is the problem for theory selection, and later refinements of this approach did not address the problem because, of course, they didn't see it. Let me re-emphasize the necessarily limited discussion of all forms of pragmatism which necessarily occurs in such a short space.

156 Dewey, "Logic: The Theory of Inquiry", pp. 105 and 107.

I shall first consider various ways of characterizing the ends or goals. Then I shall consider various ways of characterizing the means. I shall show that all characterizations of both means and ends requires a description which employs counterfactuals.

Suppose we select our theories according to our goals or ends. By that we mean that the theory is intended to serve a purpose. Some description of that purpose is required and some description of what will count as satisfying that purpose is required.

Consider, for example, van Fraassen's account of the 'aim' of science: "Scientists aim to discover facts about the world - about the regularities in the observable part of the world."[157] Such a general statement of the 'aim' will not do, as a short inquiry will determine.

If the aim is to 'discover regularities' then the obvious question is, "what is a regularity"? As we have seen in chapter three above such an obvious question has no obvious answer.

At the very least, a regularity is described in terms of a counterfactual expressed A->B. But to count as a regularity a series of counterpart counterfactuals must also be shown: A->B, B->A, -A->-B, and -B->-A.

Not all aims will be expressed in counterfactual form but the criteria for the fullfillment of such aims will always be expressed in

---

[157] Van Fraassen, p. 73.

conditional form. Therefore any description of an aim will require at least in part a description in terms of some counterfactual.

We may now proceed with this general argument to specific instances of descriptions of different aims proposed in different forms of the world view theory.

Rescher distinguishes two types of aims: first, the "cognitive or theoretical dimension of our concern for the purely intellectual aspect of information or knowledge"; and second, "the practical and affective aspects of man as agent."[158]

The second aspect quite obviously entails a description in terms of counterfactuals. We need the following sorts of descriptions: a listing of desires, needs, and wants; a description of man's physiological state; a description of actions; and a description of the consequences of such actions (relative to desires and physiological state).

I will not detail for all such descriptions the counterfactuals required. Some of those (such as desires) have been considered above. The crucial counterfactual in all such descriptions will be:

[p14] If (being of type X) performs action A then the consequences (relative to X) will be C.

A secondary counterfactual will be:

---

158 Rescher, "The Primacy of Practise", p. 4.

[p15] If, for X, (A->C), then X's (desires, etc) will be satisfied.

I shall now show that theoretical or cognitive purposes similarly require some description in terms of counterfactuals.

The method is the same as for practical purposes: first, some description of the goal will be required which may involve the use of counterfactuals; and second, some description of what counts as the description of that goal will be required.

The quick argument is this. Cognitive purposes are expressed by some 'desire' or other cognitive state. As discussed in section 5 above, the description of a desire and the description of how it is satisfied both require some counterfactuals.

The range of such cognitive purposes is much wider than might otherwise be suspected. It includes all of the 'epistemic' virtues such as simplicity and coherence.

Let us use 'simplicity' as an example.[159] We pose the question, what is it to say one theory is "simpler" than another? There is no clear answer.

Suppose we have two equally valid ways of representing a function: x=y+2 and x=sin(y). Which is simpler? It is tempting to say that x=y+2 is simpler since it is a straight line on a graph. But x=sin(y)

---

[159] It is employed by numerous theorists including Harman and van Fraassen.

might be considered simpler since there are fewer terms in the equation.

Obviously we need a rule for determining some criteria for simplicity. What might such a rule look like? Suppose we have two equally valid rules. Do we choose the simplest? But the simplest rule for choosing simple equations might actually choose more complex equations.

A clear definition of simplicity is required, a definition which must work its way through these complex objections. For each question posed in the preceding two paragraphs the definition will have to express in counterfactual form some selection of one form, and not another, of simplicity.

I shall now consider various ways of describing the means to some previously defined goal or ends.

We first begin by dividing the subject into two separate areas of investigation: the description of the means employed, and the description of the status of some means as means to some end.

Let me first consider the description of means.

In section 4 I considered the definition of pragmatism as 'practice' and the arguments stated there apply here. There is therefore no need to repeat them.

Even if the description of means does not itself require a description in terms of counterfactuals the description of means as means to an end does. We shall now consider this argument.

By stating that some means is a means to some ends we give to that means some status such as "useful" or "working". Attributing some means with this status will require a counterfactuals.

There are two ways of suggesting that some means are means to an end. First, we could argue that such means are the only possible means to that end. Second, we could argue that such means are the best means relative to some other means to that end.

Let me consider each in turn.

There are many formulations of the first option. These commonly have the form, "If anything will work, theory X will work". The first explicit formulation of this principle is found in Hans Reichenbach's discussion of induction and causation.

In short form, Reichenbach's argument is this: if our goal is to discover regularities in the world, then we must assume the hypothesis that the world is regular.

Our use of this method of inquiry (as opposed to, say, a listing of all the entities in the world) is vindicated by the success of its application.[160]

The strength of Reichenbach's argument is that no other conceivable alternative could work. Suppose, for example, we chose to predict the future by gazing into a crystal ball. This method - gazing - would be vindicated by the success of its past applications; that is, it would be vindicated inductively.[161]

It is arguable that the only means of vindicating some selection of some theory will be by the employment of some inductive hypothesis. If the only method of vindicating some means is inductively then the relevant counterfactual is easily generated. The counterfactual is:

[p16] If it worked (or was useful) in the past it will work in the future.

Having established this argument for the case of 'only possible theory' we can easily establish an analogous argument for 'best of the alternatives'. In this case the relative strength of two counterfactuals (one for each theory) is assessed and a decision made between them. Such a decision need not be expressed counterfactually

---

160 See Reichenbach, 'On The Justification of Induction', p. 161ff. The principle is also stated in Reichenbach, 'The Principle of Causality and the Possibility of its Empirical Confirmation' in "Modern Philosophy of Science". See also Feigl, 'De Principiis Non Disputandum...?'.

161 Reichenbach, 'On The Justification of Induction'.

(although it often is) but the description of the inductive strength of each theory's success rate inevitably requires a counterfactual description.[162]

I will now summarize.

I began by discussing Dewey's theory of pragmatism and distinguished between 'means' and 'ends'. For each of 'means' and 'ends' I showed that some counterfactual description is required if 'usefulness' is to be employed as a means of theory selection.

I discussed ends first. I noted that both the ends themselves and the criteria for the satisfaction of such ends must be described. I gave an example. Then I distinguished between two types of ends, the cognitive and the practical. I showed that a description of each will require a description in terms of counterfactuals.

I discussed means second. I distinguished between describing the means themselves, which was discussed in section 4, and the description of means as 'means to an end'. I suggested two ways to show that some means is a means to an ends: the means is the only means, and it is the best means. Focusing on the first, I showed how such means are vindicated inductively; this vindication requires a

---

[162] Most such arguments focus on the vindication of some or another theory in contrast to the failure of validating such a theory. Consequently little or no attention is paid to the form of the vindication. Yet it is in the form, and not the relative epistemological strength, where the trouble lies.

counterfactual description. I then showed that the best of alternative means is vindicated also.

I conclude therefore that any manner of characterizing pragmatism in terms of usefulness toward some goal or end is going to require a description using some counterfactual.

7. PROBLEMS FOR THEORY SELECTION BY CONTEXT

In the preceding four sections I demonstrated that each definition of theory selection by context which employs a theory of pragmatics or pragmatism must also contain some description which includes some counterfactuals.

In this section I shall apply the results of the preceding four sections in conjunction with the method described in section 2 and generate the objection outlined in chapter one, section 11.

First I shall briefly outline the method as described in section 2. Second I outline the counterfactuals which are required by the various pragmatic and pragmatist theories. And third I apply the argument outlined in chapter one to the method of theory selection by context.

In section 2 I showed how theory selection is guided by context. More precisely, theory selection is guided by a list of salient factors; these factors are determined by the context. The context determines such factors by identifying and quantifying variables. The

exact nature of these variables is determined using theories of pragmatics or pragmatism.

I suggested in section 2 that this procedure is exactly analogous to the model structure as described in section 1. We might say that theories of pragmatics or pragmatism act as bridge theories as described in chapter four; that is to say, they act as accessibility relations as described in chapter one.

If the analogy is as I have suggested then the objection described in chapter one should fit the use of pragmatic and pragmatist theories precisely. I argue that it does.

In preparation for this application I have at length detailed the counterfactuals which will form part of a description which (as I argued in chapter one) any accessibility relation must employ.

Allow me now to summarize that list of counterfactuals.

One or more of the following counterfactual conditionals will have to be satisfied by any use of a pragmatic or pragmatist theory of theory selection:

1. Conditions under which we say that some person accepts some theory.

2. Causal conditionals between some linguistic stimulus and some behaviour.

3. Conditional descriptions of some actions by members of some society in given circumstances.

4. Counterfactual descriptions of some object language.

5. Counterfactual transformation rules between some object language and some metalanguage.

6. Counterfactual descriptions of some 'facts of the matter' which imply linguistic presupposition.

7. Conditional rules of linguistic practices.

8. Conditional rules for the employment of rules of linguistic practices.

9. Conditional rules which describe skills.

10. Conditional rules for the employment of skills.

11. Conditional descriptions of theories.

12. Conditional rules for the employment of theories.

13. Conditions for membership in some community.

14. Conditional descriptions of some mental state such as belief, emotions and desire.

15. Causal conditionals between some mental state and some action.

16. Other conditionals expressing one mental state, such as belief, in terms of another mental state, such as physiological state.

17. Conditionals for assigning to some belief some status such as truth.

18. Conditions for assigning to some theory some status.

19. Conditional descriptions of some goal, for example, regularity or simplicity.

20. Conditions under which some means achieves that goal.

21. Causal conditionals such that the satisfaction of some goal is the consequence of some action.

22. Causal conditionals showing how some action satisfies some desire.

23. Conditions for the satisfaction of some 'epistemic virtue'.

24. Inductive conditions for the vindication of some theory.

25. Inductive conditions for the vindication of some theory in preference to another.

It may be noted that the use of a combination of pragmatic and pragmatist theories does not eliminate the need for such counterfactuals. If anything, it increases that need. More members from this list will have to be satisfied if more than one pragmatic theory is employed.

Let me now apply the objection. Notice that the objection is stated exactly as in chapter one, section 11.

1. Suppose we want to find the value of some proposition in some world (or context) in which the proposition is asserted. To do this we must select some model H which is closest to G. On the world view theory by H we refer to some theory; it is with such theories that we evaluate for truth or observation values (as noted in chapter four).

2. Closeness is defined by R. That is, it is in this case defined by some bridge theory which, in this chapter, is described by some theory of pragmatics or pragmatism. To find out how close some H is to some G we must determine the values of some atomic formulae or n-tuple relations in both G and H. That is, we must describe both the theory and the context in which it is to be employed.

3. In some cases this description will be in the form of a counterfactual. This has been proven at length.

4. We need to find the truth value of some counterfactual in G or H. So proved.

5. To find the value of some counterfactual in G go to the first step. If context determines theory selection then we must describe the context. But to describe the context we must describe a part of the world in which the theory is intended to be employed. But we needed to select that theory for just such a description in the first place. An obvious circularity takes place here.

6. To find the value of some counterfactual in H we need to find some further L which is closest to H. Cases such as this might include the description of some theory T as a 'simple' theory. This step amounts to the suggestion the we in turn need some theory of simplicity.

7. Closeness is defined by R. To find out how close some L is to some H we must find the truth values of some set of propositions in both L and H.

8. To find the truth values of some counterfactual in H go to step 6. Vicious circularity occurs if we try to define a theory in terms of itself. For example, suppose the claim for T is that T is simple. If simplicity is in turn defined by T then vicious circularity occurs.

9. L is either G or not G. If L is G go st step one. If simplicity is somehow determined by the world we need to describe the world. This produces circularity. If L is not G then it is some other H. Go to step 6. Infinite regress occurs. In the simplicity example we need a theory for the selection of some theory of simplicity.

I conclude therefore that any attempt to select theories on the basis of some pragmatically defined contextual factors results in exactly the same circularity as plagues all statements of the model-modal theory.

It is worth noting two things.

First, appeal to the similarity of the model of some theory to the world will not be an escape since, as I demonstrated in chapter four, section 7, exactly the same circularity and regress occurs.

Second, appeal to 'observations' in an attempt to break this circularity fails because, as noted at length in chapter four, what counts as an observation statement is defined by the theory the observation is intended to support. Further, as I also noted in chapter four, observations underdetermine theory selection and thus some appeal to similarity or pragmatic theories will be required in any case.

It is worth noting the impact of this argument. All of the world view theorists mentioned thus far in this thesis employ some form of pragmatic theory selection or another. With this one argument we have posed problems for Kuhn, Laudan, Lakatos, van Fraassen, Hanson, Polanyi, Kitcher, Rescher, Goodman, Quine, Feigl, and a host of others.

It was worth it.

Chapter Six

CONCLUSIONS: THE FALL OF THE MODEL-MODAL THEORY

1. A SUMMARY OF RESULTS

I have suggested the following. Logical Positivism has been replaced by a new theory, a theory I call the model-modal theory. This new theory employs a model structure which can be fairly precisely defined. And this theory, in the end, fails.

Positivism, I have argued, attempted to evaluate statements and propositions by direct inference from the world in which they are asserted. This evaluation was intended to occur independently of the context of assertion.

The model-modal theory rejects both theses. Statements and propositions are evaluated indirectly by the use of a model of the world in which such propositions and statements are asserted. The context of assertion is used in an important way to determine the selection of this model.

To show the distinction between the two theories I identified three major areas in which positivism failed: first, in the analysis of

counterfactuals; second, in the evaluation and confirmation of inductive and especially causal propositions; and third, in the distinction between observation and theory.

I discussed each of these failures respectively in chapters two, three and four. In each case I demonstrated that the distinction mentioned above applies.

Prior to this thesis the theories which succeeded positivism have never been identified as one single theory; I therefore had to prove this assertion. I have done so as follows.

First I identified a number of distinct theories following positivism and divided them roughly into two categories: possible worlds theories and world view theories.

Second I identified a model structure which I maintained is characteristic of such theories. It is represented by the triple (G,K,R) where G is an 'original', typically the world in which a proposition or statement is asserted, K is the set of models which may be used to represent G, and R is the relation between G and K such that one particular model, H, is selected from K as a model of G.

Third, I identified the use of the model structure in the evaluation of propositions or statements. In short, the idea is that the a proposition or statement asserted in G is true (or probable) if and only if a corresponding proposition or statement is true in the selected model H.

Fourth, I first suggested and then through the course of the thesis proved that all instances of both the world view theory and the possible worlds theory employ a model structure exactly analogous to (G,K,R) and evaluate statements and propositions in exactly the manner described.

Given that, first, all such theories are distinct from positivism by the rejection of the two theses, and second, all such theories employ exactly the model structure described, I conclude therefore that all such theories are in fact variations on a single theory, the model-modal theory.

I now turn to the problem which poses serious difficulties for the model-modal theory. I identify the core idea and the core problem of this theory as the selection of some model H as a model of G such that propositions and statements asserted in G may be evaluated with reference to the model H.

The selection of some model H from the set of models K is, I argued, intended to be accomplished by the accessibility relation R. This relation works in three ways. First, it restricts the set of possible models by defining some 'possibility relative to G'. Second, it obtains a preliminary selection by means of a selection function using the arguments (A,G) where A is some counterfactual antecedent and G is the 'original'. Third, it refines that selection by restricting selection to models selected by a list of salient factors defined by context.

Not all types of the model-modal theory employ all three modes of R. But all types of the model-modal theory employ at least one mode of R and all types of the model-modal theory use R to select some H.

The argument against the model-modal theory may now be stated in these terms. The argument is this. Any use of R to select some H to represent G for the purpose of analyzing or evaluating some statement or proposition in G will either result in vicious infinite regress or circularity since R, to make such a selection, must first refer to some true statements or propositions in G.

Consequently no selection may be achieved, and without a selection, no evaluation may ever be completed.

I proved this objection by considering the model-modal response to each of the problems which caused the fall of positivism: the analysis of counterfactuals, the evaluation of causal or inductive propositions, and the observation - theory distinction. In each case vicious regress or circularity was proven in exactly the same general terms as employed to demonstrate the objection for the model structure of chapter one.

In all cases the objection was the same: to describe G some model must be selected, but it is impossible to select a model without first describing G.

This objection applies against each of the three aspects of the relation R described above. To define 'possible relative to G' one

needs some description of G. To 'select' some H using a function with the arguments (A,G) one needs to satisfy the argument G. To select some model on the basis of some context of G one must describe that context of G.

Two major types of selection using R were identified: selection by similarity and selection by salient factors. Both the possible worlds theory and the world view theory employ both types of selection. Both types of selection were shown at length to fall to exactly the objection described. To show similarity, both G and H must be described. To select by context, both G and H must be described.

I conclude therefore that the major assertions of this thesis have been proved. One single theory succeeded positivism. This theory employs the model structure described to evaluate propositions and statements. And this theory falls victim to vicious infinite regress and circularity.

2. ON THE IMPOSSIBILITY OF THE MODEL-MODAL THEORY

In this section I want to suggest a slightly different argument from the one immediately preceding.

The argument is not simply that the model-modal theory fails but rather that the model-modal theory cannot possibly succeed.

Allow me first to state two theses which must be true for the model-modal theory to succeed.

[1] The evaluation of statements and propositions in some G depends on the selection of some model H.

[2] The selection of some model H depends on the evaluation of statements and propositions in some G.

That these two statements are in fact asserted has been proven in the discussion thus far. That they must be asserted is I think straightforward.

Consider [1]. If the truth of some proposition or statement in G does not depend on the selection of some H then the truth of the proposition or statement in G must depend on G. This is the positivist thesis which model-modal theorists all agree fails.

Consider [2]. If the selection of some model H does not depend in some way on a description of G then such a selection is entirely arbitrary. That means for any given selected H some other selection H' may be proposed. No criteria for the selection of H instead of H' could be given. In the absence of such a means of choosing between H and H' no proposition or statement in G could ever be evaluated.

We may therefore consider both [1] and [2] as necessary for any formulation of the model-modal theory.

The crucial word in both of those statements is "depends". Let me characterize a relation of dependence in a manner consistent with that described in chapter three: if A depends on B then this relation is represented by a series of counterfactuals such that instances of

A->B, B->A, and -A->-B tend to be true and instances of -B->-A tend to be false.

We may now characterize each such relation of dependence in these terms.

Let us symbolize 'the evaluation of some statements and propositions in G' as G. Let us symbolize 'the selection of some model H' as H.

Then we get:

[1a] G->H, H->G, and -H->-G are true and -G->-H is false.

[2a] G->H, H->G, and -G->-H are true and -H->-G is false.

We can see the contradiction clearly. There are two instances.

If [1] and [2] are both true then -H->-G will have to be both true and false.

If [1] and [2] are both true then -G->-H will have to be both true and false.

So obviously both [1] and [2] cannot be true. But if either [1] or [2] is not true then the model-modal theory fails. Therefore the model-modal theory must fail.

The reason, once we look at it clearly, should be obvious. Relations of dependency are asymmetric. If some A depends on some B then that B does not depend on that A.

This asymmetry carries over to the model-modal theory. If the truth of some statements or propositions in G depends on the selection of some model H then the selection of some model H cannot depend on the truth of some statements or propositions in G.

The model-modal theory appears to work because this contradiction does not become obvious in particular applications of the theory. If G is defined as a series of true statements and propositions {p1, p2,..., pn} then some H may be selected to evaluate some p1 and the selection of this H may be determined by some different pn.

But now we must evaluate pn. This requires some selection of H determined by the selection of some pn+1. And so on.

Either the set {p1, p2,..., pn} is infinite or it is not. If it is infinite then the selection of some H to evaluate some p can (and must) continue forever. If it is not infinite then the selection of some H must terminate at some pn and either pn cannot be evaluated (thus rendering the whole chain unevaluable) or the evaluation of pn must be based on the selection of some H by some previous pn-1, a circular evaluation.

We can see that this is exactly the problem of circularity and regress described throughout this thesis. Such circularity and regress is inevitable because the model-modal theory is impossible.

BIBLIOGRAPHY

ACHINSTEIN, Peter. 'The Problem of Theoretical Terms'. American Philosophical Quarterly 2:193-203, 1965.

ACKERMANN, Robert John. "Data, Instruments and Theory". Princeton University Press, 1969.

AYER, A.J. "Language, Truth and Logic". Second Edition. Dover, 1952.

----- "Logical Positivism". Free Press, 1959.

BARKER, John A. 'Relevance Logic, Classical Logic and Disjunctive Syllogism'. Philosophical Studies 27:361-376, 1975.

BERGMANN, M., et.al. "The Logic Book". Random House, 1980.

BENACERRAF, Paul. 'Mathematical Truth'. Journal of Philosophy 70:661-680, 1973. Reprinted in Benacerraf and Putnam, "Philosophy of Mathematics".

----- 'What Numbers Could Not Be'. Philosophical Review 74:47-74, 1965. Reprinted in Benacerraf and Putnam, "Philosophy of Mathematics".

BENACERRAF, Paul and PUTNAM, Hilary. "Philosophy of Mathematics: Selected Readings". Second Edition. Cambridge University Press, 1983.

BENNETT, Jonathon. 'Counterfactuals and Possible Worlds'. Canadian Journal of Philosophy 4:381-402, 1974.

BEROFSKY, Bernard. 'The Counterfactual Analysis of Causation'. Journal of Philosophy 70:568-569, 1973.

----- 'The Regularity Theory'. Nous 2:315-340, 1968.

BIGELOW, John. 'If-Then Meets Possible Worlds'. Philosophia 6:215-236, 1976.

BUTLER, R.J. "Analytical Philosophy". Basil Blackwell, 1962.

CHELLAS, Brian. "Modal Logic". Cambridge University Press, 1980.

CHISHOLM, Roderick. 'The Contrary - to - Fact Conditional'. Mind 55:289-307, 1946.

CHURCHLAND, Paul, and HOOKER, Clifford. "Images of Science". University of Chicago Press, 1985.

DANCY, Jonathon. "Introduction to Contemporary Epistemology". Basil Blackwell, 1986.

DAVIDSON, Donald. 'Causal Relations'. Journal of Philosophy 64:691-703, 1967. Reprinted in Sosa, "Cause and Conditionals", pp. 82-94.

DEWEY, John. "Logic: The Theory of Inquiry". Henry Holt, 1938.

EDWARDS, Paul, ed. "The Enclyclopedia of Philosophy". MacMillan, 1967.

EISENBERG, J.A. 'The Logical Form of Counterfactuals'. Dialogue 7:568-583, 1969.

FEIGL, Herbert. 'De Principiis Non Disputandum...? On the Meaning and Limits of Justification'. In Black, Max. "Philosophical Analysis".

FEIGL, Herbert, and MAXWELL, Grover, eds. "Minnesota Studies in the Philosophy of Science" Volume 3. University of Minnesota Press, 1962.

FEIGL, Herbert, and SELLARS, Wilfrid. "Readings in Philosophical Analysis". Appleton - Century - Croft, 1949.

FIELD, Hartry. 'Tarski's Theory of Truth'. Journal of Philosophy 69:347-375, 1972.

FINE, Kit. Review of Lewis, "Counterfactuals". Mind 84:451-458, 1975.

FODOR, Jerry. "Representations". MIT Press, 1981.

----- "The Language of Thought". Harvard University Press, 1975.

GAZDAR, Gerald. "Pragmatics: Implicature, Presupposition and Logical Form". Academic Press, 1979.

GIERE, Ronald. "Understanding Scientific Reasoning". Second Edition. Holt, Reinhart and Winston, 1979.

GOLDMAN, Alvin. 'A Causal Theory of Knowledge'. Journal of Philosophy 64:357-372, 1967.

GOODMAN, Nelson. 'The Problem of Counterfactual Conditionals'. Journal of Philosophy 44:113-128, 1947.

----- "Fact, Fiction and Forecast". Harvard University Press, 1983.

HACKING, Ian. "Scientific Revolutions". Oxford University Press, 1981.

HAHN, Hans. 'Logic, Mathematics, and Knowledge of Nature'. In Ayer, "Logical Positivism", p. 153ff.

HANSON, Norwood Russell. "Patterns of Discovery". Cambridge University Press, 1969.

HARPER, William, et.al., eds. "Ifs: Conditionals, Belief, Decision, Chance and Time". D. Reidel, 1981.

HEMPEL, Carl. 'Recent Problems of Induction'. In Colodny, R.G., ed. "Mind and Cosmos".

----- 'A Refutation of Verisimilitude'. Manuscript, 1986.

HEYTING, Arend. 'The Intuitionist Foundations of Mathematics'. In Benacerraf and Putnam, "Philosophy of Mathematics".

HOOKER, Clifford. 'Surface Dazzle, Ghostley Depths: An Exposition and Critical Evaluation of Van Fraassen's Vindication of Empiricism Against Realism'. In Churchland and Hooker, "Images of Science".

HUFF, Darrell. "How To Lie With Statistics". Norton, 1954.

HUME, David. "An Enquiry Concerning Human Understanding". Steinberg, Eric, ed. Hackett, 1977.

JACKSON, Frank. 'A Causal Theory of Counterfactuals'. Australasian Journal of Philosophy 55:3-21, 1977.

JAMES, William. "Essays In Pragmatism". Castell, A., ed. MacMillan, 1948.

----- "Principles of Psychology", Volume One. Dover, 1980.

KIM, Jaegwon. 'Causes and Counterfactuals'. Journal of Philosophy 70:570-572, 1973.

KITCHER, Philip. "The Nature of Mathematical Knowledge". Oxford University Press, 1984.

KRIPKE, Saul. "Naming and Necessity". Harvard University Press, 1982.

----- 'Semantical Analysis of Modal Logic'. Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik 9:67-96, 1963.

----- 'Semantical Considerations on Modal Logic'. Acta Philosophica Fennica 16:83-94, 1963. Reprinted in Linsky, "Reference and Modality".

----- 'Counterfactual Theories of Knowledge'. Lecture. University of Calgary, 1987.

KUHN, Thomas. "The Structure of Scientific Revolutions". In Neurath, et. al., "Foundations of the Unity of Science".

LAKATOS, Imre. 'History of Science and its Rational Reconstruction'. In Hacking, "Scientific Revolutions".

LAUDAN, Larry. 'A Problem-Solving Approach to Scientific Progress'. In Hacking, "Scientific Revolutions".

----- 'Realism Without the Real'. Philosophy of Science 51:156-162, 1984.

LENZ, John. 'The Pragmatic Justification of Induction'. In Feigl and Sellars, "Readings in Philosophical Analysis".

LEWIS, David K. "Counterfactuals". Harvard University Press, 1973.

----- "On The Plurality of Worlds". Basil Blackwell, 1986.

----- 'Causation'. Journal of Philosophy 70:556ff, 1973.

----- 'Counterfactuals and Comparative Possibility'. In Harper, et.al., "Ifs".

----- 'Counterfactual Dependence and Time's Arrow'. Nous 13:455-476, 1979.

----- 'Veridical Hallucinations and Prosthetic Vision'. Australasian Journal of Philosophy 58:239-249, 1980.

LINSKY, Leonard, ed. "Reference and Modality". Oxford University Press, 1971.

LOUX, Michael J. (ed.). "The Possible and the Actual". Cornell University Press, 1979.

LOVEJOY, Aurther O. 'The Thirteen Pragmatisms and Other Essays'. The Johns Hopkins Press, 1963.

MACKIE, J.L. 'Counterfactuals and Causal Laws'. In Butler, R.J., "Analytical Philosophy".

MARTIN, R.M. "Towards a Systematic Pragmatics". North-Holland, 1959.

MORRIS, Charles. "The Pragmatic Movement in American Philosophy". George Braziller, 1970.

----- 'Foundations of the Theory of Signs'. In Neurath, et.al., "Foundations of the Unity of Science", Volume One.

MORTON, Adam. 'If I Were a Dry Well-Made Match'. Dialogue 12:322-324, 1973.

MULVANEY, Robert and ZELTNER, Philip, eds. "Pragmatism: Its Sources and Prospects". University of South Carolina Press, 1981.

NEURATH, CARNAP and MORRIS (eds.). "Foundations of the Unity of Science". Volumes One and Two of "International Enclyclopedia of Unified Science". University of Chicago Press, 1938-1969.

NUTE, Donald. 'David Lewis and The Analysis of Counterfactuals'. Nous 10:355-361, 1976.

----- 'Counterfactuals and the Similarity of Worlds'. Journal of Philosophy 72:773-778, 1975.

PEIRCE, Charles. 'How to Make Our Ideas Clear'. In Weiner, "Charles S. Peirce: Selected Writings".

------ 'What Pragmatism Is'. In Weiner, "Charles S. Peirce: Selected Writings".

PLANTINGA, Alvin. "The Nature of Necessity". Clarendon Press, 1974.

POLANYI, Michael. "Personal Knowledge: Towards a Post-Critical Philosophy". Harper, 1964.

POLLOCK, John. 'Four Kinds of Conditionals'. American Philosophical Quarterly 12:51-59, 1975.

PUTNAM, Hilary. 'The Analytic and the Synthetic'. In Feigl and Maxwell, "Minnesota Studies in the Philosophy of Science" Volume 3. Reprinted in Rosenberg and Travis, "Readings In the Philosophy of Language".

QUINE, W.V.O. "From a Logical Point of View". Harvard University Press, 1961.

----- "Word and Object". MIT Press, 1960.

----- "Methods of Logic". Fourth Edition. Harvard University Press, 1982.

----- 'Ontological Relativity'. In "Ontological Relativity and Other Essays." Columbia University Press, 1969.

RAMSAY, F.P. "Foundations". Routledge and Kegan Paul, 1978.

REICHENBACH, Hans. "Experience and Prediction". University of Chicago Press, 1957.

----- 'On the Justification of Induction'. Journal of Philosophy 37:161ff, 1940. Reprinted in Feigl and Sellars, "Readings in Philosophical Analysis".

----- "Modern Philosophy of Science". Reichenbach, Maria, trans. Routledge and Kegan Paul, 1959.

RESCHER, Nicholas. "The Coherence Theory of Truth". Clarendon Press, 1973.

----- "Essays in Philosophical Analysis". University of Pittsburgh Press, 1969.

----- "Peirce's Philosophy of Science". University of Notre Dame Press, 1978.

----- "The Primacy of Practise". Basil Blackwell, 1973.

----- "Hypothetical Reasoning". North-Holland, 1964.

----- , ed. "Studies in Logical Theory". Oxford, 1968.

RICHARDS, Tom. 'The Worlds of David Lewis'. Australasian Journal of Philosophy 53:105-118, 1975.

RORTY, Richard. "Reason, Truth and History". Cambridge University Press, 1981.

RYLE, Gilbert. "The Concept of Mind". Hutchison, 1959.

SALMON, Wesley. "Logic". Prentice-Hall, 1963.

SCHLICK, Moritz. 'Meaning And Verification'. Philosophical Review 45:339-369, 1936. Reprinted in Feigl and Sellars, "Readings in Philosophical Analysis".

SHARPE, R.A. 'Laws, Coincidences, Counterfactuals and Counteridenticals'. Mind 80:572-582, 1971.

SKYRMS, Brian. 'Contraposition of the Conditional'. Philosophical Studies 26:145-147, 1974.

----- "Choice and Chance", Second Edition. Dickenson, 1975.

SOSA, Ernest, ed. "Causation and Conditionals". Oxford University Press, 1975.

STALNAKER, Robert. 'A Theory of Conditionals'. In Rescher, N., "Studies in Logical Theory". Reprinted in Harper, et.al., "Ifs".

----- "Inquiry". MIT Press, 1984.

----- 'A Defense of Conditional Excluded Middle'. In Harper, et.al., "Ifs".

STALNAKER, Robert, and THOMASON, Richmond. 'A Semantical Analysis of Conditional Logic'. Theoria 36:23-42, 1970.

SUPPE, Frederick, ed. "The Structure of Scientific Theories". University of Illinois Press, 1977.

TARSKI, A. 'The Semantic Conception of Truth'. In Feigl and Sellars, "Readings in Philosophical Analysis".

TAYLOR, Richard. 'The Metaphysics of Causation'. In Sosa, "Causation and Conditionals". Reprinted from Taylor, "Action and Purpose". Prentice-Hall, 1966.

THAYER, H.S. 'Pragmatics: A Reinterpretation of the Origins and Consequences'. In Mulvaney and Zeltner, "Pragmatism: Its Sources and Prospects".

THYGE, Per. 'On Subjunctive If-Then Statements'. Danish Yearbook of Philosophy 5:47-59, 1968.

TICHY, Pavel. 'A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals'. Philosophical Studies 29:271-273, 1976.

VAN FRAASSEN, Bas C. "The Scientific Image". Clarendon Press, 1980.

----- 'Empiricism in the Philosophy of Science'. In Churchland and Hooker, "Images of Science".

----- 'Hidden Variables in Conditional Logic'. Theoria 40:176-190, 1974.

VAN HEIJENOORT, Jean. "From Frege to Godel: A Source Book in Mathematical Logic". Cambridge University Press, 1967.

VON WRIGHT, G.H. 'On Conditionals'. In Logical Studies, pp. 127-165, 1957.

WEINER, Philip. "Charles S. Peirce: Selected Writings". Dover, 1956.

WILLIAMSON, Colwyn. 'Analysing Counterfactuals'. Dialogue 8:310-314, 1970.