15120

### CONVENTION AND THE INTENSIONAL CONCEPTS

by

### ROBERT FRANCIS HADLEY

B.A. with High Honors, University of Virginia, 1967

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

in the Department of Philosophy

We accept this thesis as conforming to the required standard

## THE UNIVERSITY OF BRITISH COLUMBIA April, 1973

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Philosophy

The University of British Columbia Vancouver 8, Canada

Date MAN 3, 1973

#### ABSTRACT

The central theme of this thesis is that our use of language is guided by linguistic conventions or linguistic rules. Substantial arguments are produced to show that we must suppose language use to be guided by such conventions. Further arguments are produced to show that the theory that language use is convention-guided can explain many facts which have not yet received satisfactory explanations. Some of the main explanatory advantages of the convention-guidedness theory are:

1) it explains the analytic-synthetic distinction.

- 2) it enables us to state, with fair precision, exactly what a concept analysis is, and how it is possible for people to use concepts without knowing the analysis of those concepts.
- it explains why people's intuitions about meaning and synonomy agree to such a large extent.
- it explains how linguistic descriptions of experience are justified by experience.
- 5) it explains why all the objects denoted by a given term often share some set of properties which can, fairly easily, be described.

One problem which has plagued earlier theories of linguistic convention has been the lack of any acceptable principle of individuation for linguistic rules. In this thesis a satisfactory principle of individuation for linguistic rules is developed. Similiarities are noted between the way computer behaviour is guided by a program and human linguistic behaviour is guided by linguistic rules. It is noted that very precise criteria exist for the identity of computer programs, and these criteria suggest criteria for the identity of linguistic rules. Other questions investigated in this thesis are:

- a) Whether Quine is right in thinking that absolutely any of our beliefs might be abandoned in the face of experiences which conflict with an accepted theory. It is concluded that this doctrine of Quine's is mistaken.
- b) Whether the notion of "logical constant" can be elucidated with any precision. This question is answered negatively. It is shown that we must take the general concepts of necessity and validity as fundamental building blocks in intellectual inquiry.
- c) Whether all necessary truths are analytic. It is shown that the answer to this question is to a large extent arbitrary.
- d) Whether necessary truths are the result of linguistic convention. It is shown that necessary truths are not, in any interesting sense, the result of convention.

### CONTENTS

Chapter 1	Introduction	1
Chapter 2	General Discussion	7
Chapter 3	Conceptual Revision	36
Chapter 4	Concept Analysis and the Analytic-Synthetic Distinction	46
Chapter 5	Recalcitrant Experiences	70
Chapter 6	Analytic Truth and Necessary Truth	87
Chapter 7	Convention and Necessary Truth	103
Bibliography		133
Footnote Index		135

#### ACKNOWLEDGEMENT

I wish to thank my thesis supervisor, Mr. Jonathan Bennett, for his invaluable suggestions and criticisms in the preparation of this thesis. Many of the ideas contained herein resulted from our discussions.

I would also like to thank Mr. Howard Jackson and Mr. John Stewart for their helpful suggestions in the preparation of the final draft.

Finally, I would like to thank the Canada Council and the University of British Columbia for Financial Assistance during the last four years.

#### Introduction

(a) "A rose by any other name would smell as sweet". In this line Shakespeare is pointing to a certain linguistic arbitrariness about the words we use. He is also pointing to a certain non-arbitrary extralinguistic fact, namely, how a rose smells. The truth of the sentence "Roses smell sweet" depends upon both of these factors.

Quine makes a related point when he says,

"It is obvious that truth in general depends on both language and extralinguistic fact. The statement 'Brutus killed Ceasar' would be false if the world had been different in certain ways, but it would also be false if the word 'killed' happened rather to have the sense of 'begat'. Thus one is tempted to suppose in general that the truth of a statement is somehow analyzable into a linguistic component and a factual component."<sup>1</sup>

But Quine also says, "My present suggestion is that it is nonsense, and the root of much nonsense, to speak of a linguistic component, and a factual component in the truth of <u>any</u> individual statement."<sup>2</sup> Quine thinks these two claims are consistent. However, on a certain interpretation, his remarks are obviously inconsistent. Quine himself provides us with a sentence whose truth value would be different if either (a) the world had been different or (b) the word 'killed' happened to have the sense of 'begat'. It is in a real sense arbitrary that the word 'killed' does not have the sense of 'begat', and it is in a real sense not arbitrary that certain things happened in the world.

It is clear that the truth value of the sentence "Brutus killed Ceasar" depends upon an arbitrary element, i.e., how we use bits of language, and a non-arbitrary element, i.e., an extralinguistic event. This certainly <u>seems</u> to refute Quine's claim that it is nonsense to speak of a linguistic component and a factual component in the truth of any individual statement.

Part of the motivation for Quine's claim is that he thinks very little can be said about the linguistic component, i.e., the meaning, which affects the truth of a sentence. This is clearly his position in "Two Dogmas of Empiricism". However, we have already said something about the linguistic component - that it is arbitrary in a particular way. In the chapters which follow I shall say a lot more about this linguistic component. I will argue that the arbitrariness we have noted is a sure sign of linguistic convention. In part I shall be leaning on an analysis of convention given by David Lewis in his book <u>Convention</u>.<sup>3</sup> But I will develop and apply the theory that our use of language is convention-guided in ways which Lewis does not consider.

(b) In addition to the notion of linguistic convention, this thesis will deal with a number of problems surrounding the so-called intensional concepts. Some of these concepts are expressed by 'analytic', 'synonomy', 'necessity', 'possibility', 'meaning', 'concept', and perhaps 'belief' and 'thought'. In "Two Dogmas of Empiricism" Quine discusses problems which pertain to the clarification of these concepts. I will now briefly summarize some of his main points. Later when it is appropriate, I will describe his arguments in more detail.

Quine focuses on the problem of clarifying our use of the terms 'analytic' and 'synthetic'. He notes that analytic statements fall into two classes. The first class consists of formal logical truths, e.g., "No unmarried man is married".

Quine characterizes a logical truth as one "which remains true under all reinterpretations of its components other than logical particles" where "a prior inventory of logical particles, comprising 'no', 'un', 'not', 'if', 'then', 'and', is presupposed."<sup>4</sup>

The second class of analytic truths is typified by

"(2) No bachelor is married. The characteristic of such a statement is that it can be turned into a logical truth by putting synonyms for synonyms; thus (2) can be turned into (a formal logical truth) by putting 'unmarried man' for its synonym 'bachelor'."<sup>5</sup>

The problem with this characterization of 'analytic' is that it appeals to the notion of "synonomy" which, Quine maintains, "is no less in need of clarification than analyticity itself."<sup>6</sup>

Quine next investigates whether our notion of "synonomy" can be adequately clarified. He notes that we can safely identify synonomy with sameness of meaning, but he questions whether we can make sense of sameness of meaning. In Quine's view we cannot identify sameness of meaning with mere coextensiveness of terms (two terms are coextensive if and only if they apply to the same objects). He says, "The general terms 'creature with a heart' and 'creature with a kidney' are perhaps alike in extension, but unlike in meaning".<sup>7</sup> Quine eventually concludes, after rejecting various attempts at clarification, that our notion of sameness of meaning cannot be adequately clarified and is unintelligible. This is a paradoxical result, given that he correctly notes that sameness of meaning cannot be identified with coextensiveness of terms. For on the one hand we find Quine making a true assertion containing the expression 'unlike in meaning', and on the other hand he concludes that the expression 'unlike in meaning' does not have a

clear use. This paradox is the focal point of an attack by Grice and Strawson which I will discuss in the following chapter.

One attempted clarification of "synonomy" which Quine discusses appeals to the notion of a definition. But this "clarification" is rejected because it places the cart before the horse. Definitions merely report what we take to be existing synonomies. Certainly we must already have the idea of synonomy before we can set about reporting a synonomy. Consequently, the notion of a definition will not help us to understand synonomy.

Another attempted clarification of "synonomy" which Quine rejects is the following: "Two expressions are synonomous if and only if they may be interchanged in all sentences in which they occur without affecting the truth value of those sentences." The problem with this "clarification", Quine thinks, is that unless the language under consideration is rich enough to contain modal adjectives like 'necessarily', it is just false that interchangeability <u>salva veritate</u> guarantees synonomy. Mere coextensiveness of terms will guarantee interchangeability in an extensional language (one which contains no intensional terms), and we know that coextensive terms need not be synonomous. So unless the language under consideration contains intensional terms like 'necessarily', the proposed criterion of synonomy is just false. However, if we modify the proposed criterion in such a way that it requires interchangeability in <u>necessity</u> contexts as a condition of synonomy, then we appeal to the idea of necessity, which is just as problematical, Quine thinks, as the idea of synonomy and analyticity.

(Quine overlooks a synonomy criterion which Frege has proposed, namely, interchangeability in <u>belief</u> contexts. This is a serious oversight, for the reason that this criterion has many advocates. I will discuss this

criterion in chapter 4.)

By now a pattern may be seen in Quine's investigation of analyticity and synonomy. His view is that, with the exception of one class of analytic sentences, i.e., the formal logical truths, our concept of an analytic sentence can only be elucidated in terms of other intensional concepts, which are just as unclear and dubious as analyticity itself. In a later chapter on Necessity I will show that even the class of formal logical truths, which Quine holds up as a paradigm of clarity, can only be identified if we presuppose an understanding of either informal validity or necessity (although Quine places both of these concepts in a class with the intensional concepts which he finds so unclear).

(c) I have, so far, summarized only a small portion of Quine's article. However, this brief summary suffices to set the stage for a discussion of Grice and Strawson's paper, "In Defense of a Dogma", which is an interesting reply to Quine's paper. I turn now to discuss Grice and Strawson's paper, and, in so doing, to discuss further aspects of "Two Dogmas".

### Chapter 1

- W.V. Quine, From a Logical Point of View, Cambridge, 1953, p.36.
  <u>Ibid</u>., p. 42.
  David Lewis, <u>Convention</u>, Cambridge, 1969.
- <sup>4</sup> Quine, <u>Op. Cit</u>., pp. 22-3.
- <sup>5</sup> <u>Ibid</u>., p. 23.
- <sup>6</sup> <u>Ibid</u>., p. 23.
- <sup>7</sup> <u>Ibid</u>., p. 21.

### Chapter 2

In this chapter I will discuss aspects of articles written by Grice and Strawson, Jonathan Bennett, and Hilary Putnam on the Analytic-Synthetic distinction. I have chosen these articles for discussion because they bring up many of the issues which I will discuss in later chapters. I. Some Remarks on "In Defense of a Dogma"

In this section I will discuss some of the arguments which Grice and Strawson bring against Quine's "Two Dogmas of Empiricism". It will become apparent that although I largely agree with Grice and Strawson, I think that their arguments are mainly suggestive and do not suffice to refute a sophisticated Quinean. In later chapters I attempt to follow up the suggestive arguments of Grice and Strawson with more powerful counter-Quinean arguments. But let us now consider some of Grice and Strawson's arguments.

1. Grice and Strawson take Quine to be denying the existence of the analytic/synthetic distinction altogether. This is natural since Quine says, "That there is such a distinction to be drawn at all, is an unempirical dogma of empiricists, a metaphysical article of faith."<sup>1</sup> Grice and Strawson strongly object to this view. They point out that there is a community of philosophers who

> "apply the term 'analytic' to more or less the same cases, withold it from more or less the same cases, and hesitate over more or less the same cases. This agreement extends not only to cases which they have been taught so to characterize, but to new cases. In short, 'analytic' and 'synthetic' have a more or less established philosophical use; and this seems to suggest that it is absurd, even senseless, to say that there is no such distinction. For, in general, if a pair of contrasting expressions are habitually and generally used in application to the same cases, where these cases do not form a closed list, this is a sufficient condition for saying that there are kinds of cases to which the expressions apply; and nothing more is needed for them to mark a distinction."<sup>2</sup>

Now Quine admits in "Two Dogmas" that the analytic-synthetic distinction could be drawn in terms of another distinction - that between expressions which mean the same and expressions which mean different things,

but he believes that this latter distinction is also suspect. In fact he doubts that we can attach any meaning to the claim that two expressions mean the same, apart from saying that the two expressions are coextensive. This position of Quine's runs counter to the very ordinary belief that whereas 'oxygen' and 'a gas produced by plants during photosynthesis' may be coextensive, they certainly do not mean the same.

Grice and Strawson point out that the ordinary language expression 'means the same as' has an established use, just as the philosopher's technical term 'analytic' does. By and large people agree about which expressions mean the same, about which mean different things, and they hesitate over approximately the same group of expressions. This fact is sufficient to establish that there is some distinction between expressions which are thought to mean the same and expressions which do not.

Furthermore, it certainly makes sense to say that term x has a meaning and that term y has a meaning. From this it is quite plausible to infer that it makes sense to say that x and y have the same meaning. So if Quine wishes to deny that there is <u>any</u> distinction marked by the expressions 'means the same' and 'means something different', then he is probably wrong. And by Quine's own admission, if there is a distinction between synonomous and non-synonomous expressions there is also a distinction between analytic and synthetic sentences.

In an article entitled "Quine on Meaning and Existence" Gilbert Harman objects to the line of argument I have been describing. He argues as follows. The fact that many philosophers agree about what sentences are analytic and synthetic (and about which expressions mean the same) does not

entitle us to infer that some sentences actually are analytic, any more than the fact that people once agreed in their use of the terms 'witch' and 'nonwitch' entitles us to infer that some things are witches. In Harman's view, to claim that a sentence is analytic is to make an explanatory claim about how the sentence comes to be true. If no sentence fits this explanatory claim, then no sentences are analytic. Harman does concede, however, that

> "There is a distinction between truths that seem to be analytic and truths that seem synthetic, and that distinction "underlies" general agreement on the use of 'analytic' and 'synthetic' with respect to an open class."<sup>3</sup>

Harman also concedes that there is a distinction between people who seem to be witches and people who do not.

Now given these concessions it is difficult to see that Harman has raised any serious problem for Grice and Strawson's line of reasoning. For the interesting point which Grice and Strawson make still remains; namely, that there must be <u>some</u> distinction between those sentences which appear to be analytic and those which appear to be synthetic. Whether it is proper to describe this distinction as the A-S distinction is a question worth investigating. But regardless of how this question is answered there still remains the problem of explaining why certain sentences appear, to large numbers of people, to be analytic, while others do not so appear. Also, if there is to be any point to saying that some sentences <u>appear</u> to be analytic, but no sentences <u>really</u> are analytic, then we must be able to say in what way sentences which appear to be analytic fail to be analytic. And we should be able to say what it is about seemingly analytic sentences which makes them appear to be analytic. In order to be able to do both these things we would need to have a fairly clear account of what would count as an analytic

sentence. (Just as we can give a fairly clear account of what would count as a witch.) It is also interesting to note that if we were in a position to say in what ways seemingly analytic sentences resemble real analytic sentences, we would also be able to say what distinguishes sentences which seem to be analytic from those which seem to be synthetic.

2. It was mentioned in the introduction that Quine has noticed that 'analytic' has usually been defined in terms of a family of (allegedly) interdefinable intensional terms, e.g., 'necessity', 'impossibility', and 'synonomy'. Quine also believes that the family of interdefinable intensional terms is <u>closed</u> insofar as no member of the family can be formally defined except in terms of other members of the family. Consequently, if one wishes to understand the intensional term, 'analytic', one can only turn to other intensional terms, which Quine claims, stand in as much need of clarification as 'analytic' itself.

In their defense of the analytic-synthetic distinction Grice and Strawson are willing to concede, for the sake of argument, that 'analytic' may belong to a closed circle of mutually interdefinable intensional concepts, but they deny that this fact constitutes any reason for doubting the intelligibility of 'analytic'. It may be equally true of the term 'truth', that it belongs to a closed circle of interdefinable terms, e.g., 'false', 'statement', 'fact', 'denial', 'assertion', etc., but this does not lead us to doubt the intelligibility of the term 'truth' (a term which Quine contrasts favorably with 'analytic'). Grice and Strawson conclude that Quine must have some special reason for being suspicious of 'analytic', apart from the fact that it is allegedly definable only in terms of intensional concepts. This con-

clusion is a safe one, given that Quine constantly remarks that 'analytic' can only be elucidated in terms of notions which are just as unclear as 'analytic' itself. The question arises, why does Quine think that the family of intensional concepts is unclear in some way that extensional concepts like 'truth' are not. Quine might say that he just does not understand intensional terms like, 'analytic', 'synonymous'. This would be a paradoxical claim for Quine to make, given that he also claims that 'analytic', 'synonymous', 'necessary', and 'inconsistent' are interdefinable terms. For in claiming that, e.g., 'necessity' and 'impossibility' are <u>interdefinable</u>, he tacitly admits that he understands the notion of synonomy and that philosophers do use these terms in a fairly clear and systematic way.

Perhaps Quine's difficulty with the circle of intensional terms is expressed when he says, "I do not know whether the statement "Everything green is extended" is analytic. Now does my indecision over this example really betray an incomplete understanding, an incomplete grasp, of the meanings of "green" and "extended"? I think not. The trouble is not with "green" or "extended", but with "analytic"."<sup>4</sup>

But if Quine's doubts about the intensional family of terms stem from his hesitation over whether to say "Everything green is extended", and related cases, are analytic, then his doubts are ill founded. Grice and Strawson point out that whatever hesitation exists about saying that "Everything green is extended" is analytic also exists when we replace 'analytic' with 'true'. But Quine would not conclude that our concept of truth is hopelessly vague just because we cannot always say whether a sentence is true. Grice and Strawson conclude that Quine does not produce convincing

reasons for doubting the intelligibility of the intensional concepts. I agree with Grice and Strawson, but I sympathize with Quine insofar as he seeks a theoretically satisfying account of "analytic".

3. It was mentioned earlier that Grice and Strawson admit that intensional terms may be formally definable only in terms of other intensional terms, but they maintain that intensional terms can also be <u>informally</u> explained in non-intensional terms. If they are right, then it is possible to break out of the circle of intensional terms, in an informal way. Let us consider Grice and Strawson's example of such an informal explanation.

They contrast two sentences:

A. My three year old child understands Spinoza.

B. My three year old child is nine years old.

They point out that if anyone asserted A we would be very skeptical, but would understand the assertion. However, if someone asserted B we would be unable to make sense of that person's assertion, we would be bewildered. Statements which are bewildering in the way that B is bewildering are inconsistent. Statements like A are difficult to believe, but they are consistent. Grice and Strawson conclude that "The distinction in which we ultimately come to rest is that between not believing something and not understanding something; it would be rash to maintain that <u>this</u> distinction does not need clarification, but it would be absurd to maintain that it does not exist."<sup>5</sup>

Now someone might object that Grice and Strawson have not really broken out of the circle of intensional terms. One might claim that 'understand' is an intensional term. But it would be absurd for someone to

that they do not <u>understand</u> what it is to understand something, and as Grice and Strawson point out it would be absurd to maintain that no distinction exists between not believing something and not understanding what it would be for a particular sentence to be true. So this objection does not seem to have much force.

Jonathan Bennett raises a different objection. In "Analytic-Synthetic" he points out that while it is true that a statement like (C) "It is raining and also not raining" is bewildering, it is also true that statements that are normally considered synthetic are bewildering when they are asserted in certain situations. For example, if a friend is staring out the window at the sunshine and asserts (D) "It is raining so hard that you shouldn't go out" his remark is equally bewildering as (C). So it appears that we cannot distinguish inconsistent sentences by their bewildering property.

Grice has tried to answer this objection. He claims that we may distinguish the former sentence from the latter by the fact that the former sentence is <u>prima-facie</u> bewildering, whereas the latter is bewildering only in some circumstances. Bennett rejects this answer on the grounds that the notion of "<u>prima-facie</u> bewilderment" can only be elucidated by some such notion as "a sentence's being bewildering to anyone who knows the <u>meaning</u> of its constituent terms even if that is all he knows",<sup>6</sup> and this raises all the problems about meaning which one wishes to avoid.

I think it can be shown that Bennett's objection does not have much force, however. For suppose we distinguish the two cases of bewilderment as follows: "The only story we have to tell about C, for its assertion

to be bewildering, is that the expressions in C bear their usual meanings. The assertion of D, on the other hand, is bewildering only if the expressions in D bear their usual meanings <u>and</u> we believe its asserter believes it to be false. What is bewildering in this latter case is that the <u>asserter</u> seems to have contradictory beliefs.

It is true that this way of drawing the distinction appeals to the idea of a sentence bearing its usual meaning, but only in a harmless way. Whenever a person advances a philosophical theory there is a background assumption that the sentences constituting the theory bear their usual meanings. And whenever we judge the <u>truth</u> of a sentence we assume that the sentence bears its usual meaning. It can hardly count against my way of clarifying the notion of <u>prima-facie</u> bewilderment that it makes explicit an assumption that is always present in any discourse. The fact that we must understand a sentence before we can tell whether it is bewildering does not count against the bewilderment/non-bewilderment distinction any more than it counts against the true/false distinction that we must understand a sentence before we can say that it is true.

Perhaps this answers Bennett's objection, but a feeling of dissatisfaction may yet remain. One feels reluctant to say that our concept of inconsistency is adequately clarified by the distinction between not believing an assertion and not understanding an assertion (or being bewildered by an assertion), because this distinction is not theoretically deep. Grice and Strawson concede this when they say "it would be rash to maintain that <u>this</u> distinction does not need clarification." I will argue in a later chapter that our concept of inconsistency (and its cousin, necessity) cannot be clarified in a theoretically deep way. Consequently our concept of formal

logical truth cannot be clarified in a theoretically deep way, either (or so I will argue). But I also hope to show that the analytic-synthetic distinction can be clarified in a theoretically satisfying way, even though we must take the concept of logical truth as one of our building blocks.

In "Two Dogmas" Quine briefly discusses the kind of synonomy which results from stipulative definition. Concerning it he says, "Here the definiendum becomes synonymous with the definiens simply because it has been created expressly for the purpose of being synonymous with the definiens. Here we have a really transparent case of synonymy created by definition; would that all species of synonymy were as intelligible."<sup>7</sup> An example of the kind of stipulative definition Quine is referring to is the convention of writing 'B.A.' for bachelor of arts. Quine finds this kind of synonymy acceptable and intelligible, but complains that he cannot make sense of other kinds of synonymy.

Grice and Strawson point out that Quine's position is like that of one who says "I can understand what it means to say that one thing fits into another, or that two things fit together, in the case where one was specially made to fit the other; but I cannot understand what it means to say this in any other case."<sup>8</sup> Grice and Strawson contend that Quine's view is incoherent on this point. We can only understand stipulative definition because we know what it is to follow a linguistic practice over a period of time. And if we know what it is to follow a linguistic practice over time, we should be able to understand that the same linguistic practice can govern the use of two different words. That, according to Grice and Strawson, amounts to understanding synonymy.

I think Grice and Strawson are on the right track here, but much more needs to be said. For the sophisticated Quinean might object that we need to be able to distinguish regularities which are due to linguistic custom from regularities which are due to natural law. For example, why do we say that the coextensiveness of the expressions 'bachelor' and 'unmarried adult male' is due to linguistic practice, but the coextensiveness of 'oxygen' with 'gas which plants release during photosynthesis' is due to a natural regularity. If we are not able to explain the distinction between coextensiveness which is due to convention and that which is due to natural law must we not conclude that there is no such distinction?

I will not try to answer this question. Rather I will attempt to show, in a later chapter, how we can draw the required distinction. In so doing I will say a lot more about linguistic conventions and rules, and attempt to give a much firmer foundation to the position Grice and Strawson sketch.

4. Grice and Strawson challenge Quine's doctrine that absolutely any belief might be sacrificed in order to keep a theory consistent with experience. They contend that some sentences, i.e., the analytic ones, can only be rejected if a change in meaning occurs in the sentence being rejected. In this way analytic sentences differ from synthetic sentences. According to Grice and Strawson the rejection of a synthetic sentence may reflect "a change of opinion solely as to matters of fact." But the rejection of an analytic sentence always requires a meaning change.

Once again, I agree with the position Grice and Strawson sketch, but more needs to be said. For the sophisticated Quinean may contend that

we cannot really distinguish the case where the rejection of a sentence reflects only a change in beliefs from the case where the rejection also reflects a change in meaning. He would probably insist that the idea of a meaning change is what needs clarifying.

In the following chapter I will argue that we can distinguish a meaning change from a change in beliefs. Also I will attempt to elucidate the idea of a meaning change. The conclusion I will reach entirely supports Grice and Strawson when they say,

> "The point of substance that Quine is making, by this emphasis on revisability, is that there is no absolute necessity about the adoption or use of any conceptual scheme whatever, or, more narrowly and in terms that he would reject, that there is no analytic proposition such that we <u>must</u> have linguistic forms bearing just the sense required to express that proposition. But it is one thing to admit this, and quite another thing to say that there are no necessities within any conceptual scheme we adopt or use, or, more narrowly again, that there are no linguistic forms which do express analytic propositions."<sup>9</sup>

II. Some Remarks on "Analytic-Synthetic"

1. Recall the remarks of Grice and Strawson concerning the use of 'analytic' and 'synthetic'. They point out that most philosophers agree about which sentences to call 'analytic', which to call'synthetic' and most philosophers hesitate over roughly the same cases. This fact constitu**tes** very strong evidence for saying that these words do mark <u>some</u> distinction, though, it is quite possible that many philosophers hold mistaken beliefs about the nature of this distinction.

In "Analytic-Synthetic" (reprinted in Necessary Truth, Summer

and Woods) Bennett describes a sophisticated Quinean. The sophisticated Quinean holds that 'analytic', 'synthetic' and other intensional terms mark genuine distinctions, but he desires a Quinean analysis of these distinctions. Bennett offers such an analysis. In so doing he tries to elucidate the Quinean theory that certain beliefs are deeply embedded in our conceptual scheme, while others lie near the periphery of our conceptual scheme, and still others lie somewhere in between. Bennett cashes the metaphors "deeply embedded" and "near the periphery" into the more literal notions of "highly indispensible" and "relatively dispensible". Roughly the picture he describes is as follows:

> "Accepted sentences of the form (i) 'The temperature of suchand-such a star is such-and-such' depend, for those who accept them, on sentences of the form (ii) 'Temperature correlates with light emissions in such-and-such ways', and these depend on sentences of the form (iii) 'Temperature correlates with mercurycolumn readings in such-and-such ways', and these in their turn depend on sentences along the lines of (iv) 'Temperature has to do with the obtaining of such and such sensations.' Rejection of (ii) jeopardises (i) and all that depends on it; rejection of (iii) jeopardises (i) and (ii); rejection of (iv) jeopardises all the other three."<sup>10</sup>

"When I wish to say that one sentence shares a general term with another and has more depending on it than depends on the other, I shall say that it is less dispensible than the other."<sup>11</sup>

In order of dispensibility (i) is the most dispensible and (iv) the least dispensible. A sentence is less dispensible than another if and only if the first forms part of the grounds for accepting the second, but not vice-versa. According to Bennett, if all the sentences sharing the same general term F are ranked according to dispensibility, the least dispensible sentence provides the test for Fness which is used in establishing the truth of more dispensible sentences containing F, and these in turn provide the test for Fness which is used in establishing even more dispensible sentences containing

F.

That fairly accurately describes Bennett's picture of the distinction between highly indispensible and fairly dispensible sentences. It is his view, in "Analytic-Synthetic" that the actual distinction marked by 'analytic' and 'synthetic' is this dispensibility distinction. It may be that most people who use these words would not recognize the dispensibility distinction as the distinction they intend, but, according to Bennett, this would only show that most philosophers have been mistaken about the nature of this distinction. If we look at the way people actually <u>use</u> 'analytic' and 'synthetic' we shall find that it corresponds with Bennett's use of 'highly indispensible' and 'relatively dispensible'.

According to Bennett this reconstruction of the A-S distinction has the advantage that it breaks out of the circle of intensional terms, and allows us to explain the distinction to someone who does not understand intensional terms. I am skeptical whether this is true, since Bennett's theory appeals to the fact that one sentence can form part of the ground for the acceptance of another sentence, and this involves the idea of logical consequence, which (I shall argue in a later chapter) involves the idea of necessity. I will ignore this problem at present, however, for there is a more straightforward problem to be dealt with. It is this: The sophisticated Quinean holds the view that absolutely any sentence (excluding perhaps a special class of observation sentences) may be sacrificed in the face of recalcitrant (with respect to our theories) experience, while any other sentence may be retained. This view entails that under some circumstances we might be willing to sacrifice the highly indispensible sentence (iv), and yet retain (i), (ii), and (iii). But if this is so, then how can we say

that (iv) constitutes the grounds for the acceptance of (i), (ii), and (iii)? If (iv) is more indispensible than the other three in the sense that "rejection of (iv) jeopardises all the other three", then it is not possible to reject (iv) and yet retain all the other three. Furthermore, if (iv) provides the test for temperature in virtue of which all the other three came to be established, it is difficult to see what one would be asserting when one asserted any of the other three after having rejected (iv).

The sophisticated Quinean, who wants to use Bennett's dispensibility ranking theory, might try to avoid the difficulty I describe by adopting the kind of position Putnam describes in "The Analytic and the Synthetic" when he discusses law-cluster concepts. According to Putnam many of our concepts are derived from clusters of laws. For example, the word 'temperature' does not derive its meaning from any single law, but rather from a whole cluster of laws in which it appears. Consequently, there is no single law whose rejection would destroy our concept of temperature. We could reject any particular law containing 'temperature' and yet retain our other beliefs about temperature, because our concept of temperature would remain nearly intact. It might still be true that some laws about temperature are more indispensible than others, because some laws might contribute more meaning to 'temperature' than others. It might also be true that a single cluster of laws about temperature is more indispensible than all our other beliefs about temperature, while no single law is absolutely indispensible. Thus Putnam's view seems to be compatible with the kind of ranking of dispensibility which Bennett describes.

Unfortunately, for the sophisticated Quinean, this modification will not suffice. For while it may be true that the rejection of particular

laws about temperature does not significantly alter our concept of temperature, it will still be true (assuming a dispensibility ranking)that the rejection of the whole cluster of high level laws about temperature jeopardises our low level beliefs about temperature, and so it will still be true that we cannot simultaneously reject all our high level beliefs about temperature and retain our low level beliefs intact. There is a basic incompatibility between holding that some beliefs constitute grounds for other beliefs, and also holding that any belief may be rejected while any other belief is retained.

An interesting point is emerging here. Namely that anyone who holds that each of our beliefs is open to rejection while any other belief may be retained is going to have an impossible time making sense of the idea that one belief constitutes the grounds for accepting another belief. If any conclusion can always be rejected while any set of premises can be retained, then what sense can we make of <u>logic</u> - not just traditional logic, but any logic? It will not help to say that a premise p logically entails a conclusion q provided rejection of q and retention of p would <u>entail</u> an enormously complicated revision in the remainder of our beliefs, for it is the very concept of entailment which is in question.

The upshot of all this is that if the sophisticated Quinean, or Bennett, wants to retain his dispensibility ranking theory, he must reject the Quinean doctrine that any belief can be rejected while any other belief is retained. This would eliminate one problem for the dispensibility ranking theory, but other problems remain. For example, there is still the problem that on this theory some beliefs are logically dependent upon others, and

this may well involve the intensional notion of necessity. Other problems arise from Bennett's view that the least dispensible sentence provides the test for Fness which is used in establishing the truth of more dispensible sentences containing F, and these in turn provide the test for Fness which is used in establishing even more dispensible sentences containing F.<sup>12</sup> There is a problem about deciding which sentences provide the test for Fness. And to a Quinean the view that a definite class of sentences provides the test for Fness may appear just as much a dogma as the view that the conventions guiding our use of 'F' are exact, or that 'F' can be precisely defined. In any case it seems clear that a lot more needs to be said about what it is for a class of sentences to provide the test for Fness.

I will discuss this issue at length (in a slightly different form) when I discuss the question of how conventions guide the use of particular words, why there must be such conventions, and how we can distinguish conventions from regularities. But this discussion, like the discussion of necessity, must await a later chapter.

2. In "Analytic-Synthetic" Bennett argues for the view that "any analytic sentence may become false through a meaning change which is brought about by the occurrence of recalcitrant experiences."<sup>13</sup> This accords with the position Grice and Strawson adopt in "Defense of a Dogma", but whereas Grice and Strawson stress the fact that an analytic sentence <u>cannot</u> become false <u>unless</u> a meaning change occurs, Bennett stresses the fact that an analytic sentence can become false if a meaning change occurs.

Now Bennett's claim may not seem very controversial. Almost everyone concedes that the truth value of a sentence is partly a function

of the meaning of the sentence, and there is nothing God-given about the meaning of a sentence; it can change if people's linguistic habits change. What is controversial about Bennett's claim, however, is the idea that a meaning change can be brought about by the occurrence of recalcitrant experiences. Bennett briefly describes how this might happen. He says,

> "Associated with any analytic sentence there is a range of synthetic sentences stating facts about the world in virtue of which it is convenient that the words in the analytic sentence should have the meanings they do have; suppose a falsification of a judiciously selected sub-set of these synthetic sentences, and you are well on the way to describing a state of affairs which invites the falsification of the analytic sentence."<sup>14</sup>

Now there is a certain misleadingness when Bennett says "you are well on the way to describing a state of affairs which invites the falsification of the analytic sentence." For this might easily lead us to believe that there is some state of affairs which falsifies the analytic sentence, which, in turn, would lead us to infer that an analytic sentence can assert something false.

But this is not in fact Bennett's view. Rather Bennett is claiming that a change in our factual beliefs can make it <u>convenient</u> to change the meaning of an analytic sentence in such a way that the sentence ceases to be analytic, and ceases to be true. Of course, Bennett believes that at the time a sentence is analytic it is true. This, however, makes it somewhat misleading for Bennett to claim that we may reject our belief in any analytic sentence. For it is strange to say, at  $T_2$ , that we have rejected our belief in S, when we still believe that S was true at  $T_1$ . We may have rejected some beliefs concerning those "facts about the world in virtue of which it is convenient that the words (in S) have the meaning they do have", but

those beliefs are distinct from the belief expressed by S at  $T_1$ . In general, a sentence, S, does not make an assertion about the facts which make it convenient for S to have the meaning it has. Consequently, if an analytic sentence changes meaning because some of these meaning-related beliefs have been falsified, that does not constitute a reason for saying that the belief formerly associated with the analytic sentence has been falsified.

The reader may have noticed the use of intensional terminology in the preceding discussion. Bennett introduces this intensional language when he talks about circumstances which could lead us to change the meaning and truth value of analytic sentences. However, he offers a non-intensional analysis of this intensional language when he says, "The proposition (meaning) expressed by  $S_1$  at  $t_1$  is different from the proposition (meaning) expressed by  $S_1$  at  $t_2$  if and only if an appropriate set of sentences of the form " $S_1$  is true if and only if  $S_n$  is true" which are highly indispensible at  $t_1$  are not highly indispensible at  $t_2$ . So that a hitherto analytic sentence can be denied only if it comes to express a different proposition from the one it formerly expressed."<sup>15</sup>

Now the success of this non-intensional analysis of meaning change depends upon the intelligibility of Bennett's "dispensibility" ranking theory. But we have already seen that there are problems with this theory, and a lot more needs to be said about it. Consequently, although I agree with Bennett that an analytic sentence can only be rejected if its meaning and truth value are changed, I do not think he has demonstrated this thesis, or clearly explained it in non-intensional language.

3. In the last section of "Analytic-Synthetic" Bennett discusses possible conflicts between Quine's beliefs that (a) certain experiences (the recalcitrant ones) can force us to make some revision in our total network of beliefs, and (b) that when we revise our total set of beliefs in the face of a recalcitrant experience absolutely any belief is a possible candidate for revision. I do not intend to discuss this portion of Bennett's paper in any detail, partially because Bennett now rejects the arguments contained therein. But I do wish to call attention to one of his arguments, because it relates to another argument which I will examine in a later chapter.

Bennett shows that if Quine is to retain (b) then he must hold that our total set of beliefs is infinite. He argues as follows: if our set of beliefs is finite, and if some experience occurs which is recalcitrant with respect to those beliefs, then we cannot simultaneously affirm the existence of the recalcitrant experience and the conjunction of all the sentences we call true, and so "there is a sentence (albeit a long one) which is, in isolation, strongly disconfirmed by experience."<sup>16</sup> Bennett suggests that Quine might escape this conclusion by supposing that our total set of beliefs is infinite, and Bennett goes on to investigate this possible move on Quine's part.

Rather than discuss the complex arguments which Bennett considers, however, I suggest that the issue whether our theory consists of an infinite or finite number of sentences is irrelevant. For whether our theory is finite or infinite it is clear that we cannot simultaneously affirm the existence of a recalcitrant experience and the truth of our total theory. And this means that a certain belief is forced on us by experience, (namely the belief that

some revision in our total set of beliefs is required) and its negation is strongly disconfirmed by experience. (This argument is but a sketch. It will be considerably expanded in my chapter on recalcitrant experience.)

III. Some Remarks on Putnam's "The Analytic and the Synthetic"

In this section I will discuss some aspects of Putnam's theory of the A-S distinction. Although I will reject many of Putnam's doctrines, I think much of what he says is very suggestive. In particular, his definition of 'analytic' is suggestive of a theory which I later develop. But let us now consider Putnam's theory.

According to Putnam Quine is wrong in his literal thesis that there is no analytic-synthetic distinction at all. There are statements which it would be unreasonable for anyone to hold to be false at any time and in any circumstances. These are the statements which philosophers have cited as paradigm cases of analytic statements. There are also statements which can be rejected on the basis of isolated experiments. These are statements which philosophers have cited as paradigm cases of synthetic or contingent statements. By and large people agree on what are the paradigms of analytic and synthetic statements, and (as Grice and Strawson also point out) "Where there is agreement on the use of the expressions involved with respect to an open class, there must necessarily be some kind of distinction present."<sup>17</sup> So there is an A-S distinction, but the importance of the distinction has been immensely overestimated because people have failed to

realize the following points: (a) The distinction is not an exhaustive one. Many statements are neither analytic nor synthetic. (b) Most philosophers have mistakenly taken some important and interesting statements to be analytic. Among these are the statements of logic and mathematics, and some statements of physics. Having mistakenly placed these statements in a class with such utterly trivial statements as "All bachelors are unmarried", these philosophers have assumed that the interesting statements, like the trivial statements, owe their truth to linguistic convention. (c) Since all analytic statements are <u>obviously</u> analytic and utterly trivial one cannot hope to use the A-S distinction as a tool for discovering interesting truths. Putnam puts the matter somewhat as follows:

" 'Chair' may be synonymous with 'movable seat for one with a back' but that bakes no philosophic bread and washes no philosophic windows. It is the belief that there are synonymies of a deeper nature -- synonymies and analyticities that cannot be discovered by the lexicographer or the linguist but only by the philosopher -- that is incorrect."<sup>18</sup>

Let us consider each of these three points in turn. Concerning (a), we have already seen that the A-S distinction has some borderline cases. For example, the sentence "Everything green is extended" is not clearly analytic or synthetic, but it is not clearly true or false either. The mere fact that the A-S distinction is not exhaustive should not lead us to doubt the importance of this distinction, any more than it should in the case of the true-false distinction.

Concerning (b). Putnam holds that statements of mathematics and logic belong in a class with high-level physical laws, that statements in this class are neither analytic nor synthetic, and that any statement in this indeterminate class can be rejected as false, without altering any of our concepts to a

significant extent. Although Putnam never does <u>prove</u> that mathematical and logical laws belong in this class of statements, his reasons for holding this view are fairly apparent. They stem from his theory about law-cluster concepts which I briefly described in the section on Bennett.

Remember that, according to Putnam, all mathematical, logical, and scientific terms derive their meaning from whole clusters of laws in which they occur, and none of these terms derive their meaning from single laws. For this reason he thinks that any single mathematical, logical, or physical law could be rejected, without destroying the concepts expressed by terms appearing in the rejected law. For example, the word 'and' appears in, and derives its meaning from countless logical laws and theorems. Consequently, it would be arbitrary to single out any particular logical law and claim that it completely defines the meaning of 'and'. So it would be a mistake, in Putnam's opinion, to think that any logical law is "true by definition", in the sense that it follows from the definition of 'and'. Putnam concludes that since logical laws do not hold in virtue of any definition, or simple convention, they are not necessary truths, and are open to rejection. Perhaps this is because he identifies "necessary" with "true by definition or convention". In later chapters I try to show that this identification is a mistaken one.

I think there are some serious problems with Putnam's view (some of which have already been mentioned in my discussion of the sophisticated Quinean). They are: (A) Putnam, who is a sophisticated Quinean, is going to have an extremely difficult time making sense of the idea that a conclusion can logically follow from certain premises. For since he holds that any

logical law (or rule of inference) could be rejected, he is committed to holding that, in principle, any conclusion could be rejected while any set of premises is retained. Of course, he will want to claim that this is <u>possible</u> only if revisions are made elsewhere in our total system of beliefs. That is, he will claim that the rejection of a logical law or logistic system <u>logically entails</u> a revision elsewhere in our set of beliefs. But the question arises, according to what logic does the rejection of a logical law entail some other revision? Presumably it is not some second order logic which is not open to rejection. And not every logic will entail some other revision. So what happens if the logic being rejected is the same as the logic which is supposed to entail some other revision? For example, the logic which now motivates us to make revisions in our theory is a logic in which consistency is a central requirement. Consequently, if we reject this logical system, including the logical requirement of consistency, we reject the motivation for some other revision.

I think the reason Putnam, and Quine, have not faced up to this problem is that they really believe that we are, in some absolute way, logically committed to making some other revision when we reject a logical principle; and this is just incompatible with their belief that any logical principle can be rejected.

(B) Putnam denies (i) that there is any set of logically necessary and sufficient conditions which defines scientific, mathematical, and logical terms. He also holds (ii) that these terms derive their meaning from the cluster of laws in which they occur. Now there is a problem about how to interpret (ii) in such a way that (ii) is consistent with (i). We might interpret (ii) as the claim that (iii) the meaning of law-cluster terms is
some complicated <u>function</u> of the cluster of laws in which they occur (or as the view that the use of law-cluster terms is determined in some complicated way by the way we use the whole cluster of laws in which they occur). But this interpretation of (ii) seems to be inconsistent with (i). For if the meaning of some term, T, is a function of some set of laws, L, then this function must be describable. (I assume that whatever exists can be described, at least in principle). And if the meaning function which relates T to the set of laws L can be described, then we can construct a definition of T in terms of the meaning function and L as follows:

The meaning of T = f(L)

There is a modification of Putnam's view (or perhaps it is merely a generous interpretation) which would avoid the difficulty just described. It is this: suppose we admit that the meaning of scientific terms is a complicated function of some cluster of physical laws. We also admit that this function is describable, and that, in principle, scientific terms are definable. From this it might follow that <u>some</u> general statements containing scientific terms are "true by definition" or analytic, but this is certainly compatible with holding that the particular laws, which constitute the defining cluster of laws, are always open to rejection. For example, suppose that a scientific term, T, can be defined in the following complicated way:

We apply T to those cases where most of the following laws are

satisfied:  $L_1, L_2 \dots L_n$ . Clearly the rejection of any one of the laws  $L_1 \dots L_n$  would not bar us from applying T in a particular case, (though rejection of very many of these laws would). On the other hand, not every statement in which T occurs can

be rejected without destroying the concept expressed by T.

Putnam might accept the modification just suggested. If he does he should abandon his belief that there are no deep, or non-trivial synonymies and analyticities for philosophers to discover. For there is no reason to believe that the meaning function which relates a particular law-cluster term to a whole-cluster of laws will usually be trivial or obvious.

To say that a particular law-cluster term <u>derives</u> its meaning from some identifiable set of laws is already to say something interesting about the meaning of the law-cluster term. But to describe in any detail the meaning relationship between a particular term and some cluster of laws should be a complex and interesting task. It is, of course, open to Putnam to deny that this meaning relationship, or meaning function, can be described in any detail, but this would cast considerable doubt on the intelligibility of Putnam's claim that the law-cluster terms <u>derive</u> their meaning from a whole cluster of laws.

Discussion of (c). By now some reasons have been given for doubting (c), i.e., the view that all analytic sentences are utterly trivial and uninteresting. I do not think Putnam has demonstrated that the truths of mathematics and logic are non-analytic, and they are far from trivial. In what follows I will argue that Putnam's view about the triviality of all analytic sentences is not even consistent with his own formal definition of 'analytic sentences'. Here is Putnam's definition:

> "(1) The statement has the form: "Something (Someone) is an A if and only if it (he, she) is a B," where A is a single word. (2) The statement holds without exception, and provides us with a <u>criterion</u> for something's being the sort of thing to which term A applies." <u>Criterion</u> is defined as follows: "A statement of the form "Something is an A if and only if it is a B"

provides a criterion for something's being a thing to which the term A applies if people can and do determine whether or not something is an A by <u>first</u> finding out whether it is a B."<sup>19</sup>

Very little reflection will show that a statement which is analytic according to Putnam's criteria will be obvious and trivial only if the criterion in guestion is obvious and trivial. But there is no reason, in general, to think that the criterion, by which people determine whether something is an A, is at all obvious or trivial. In the case of "Someone is a bachelor if and only if he is an unmarried adult male" it may be obvious that we determine whether someone is a bachelor by determining that he is an unmarried adult male, but we apply many of our words without consciously knowing what considerations determine our applications of the words. Perhaps this is because the procedure we follow when we apply many words is very complicated, and we internalize the procedure at an early age. For example, people can and do determine whether sentences are grammatical. The fact that people agree to a great extent about which sentences are grammatical and which are not suggests that there may be some unconscious processing of auditory (or visual) information which people first do before deciding whether a sentence is grammatical. This unconscious processing of information might occur according to a complex decision procedure which is conventionally associated with the word 'grammatical'.

Oddly enough Putnam argues for this view in an article entitled "Some Issues in the Theory of Grammar". He says, "This act of classifying sentences as grammatical or ungrammatical seems to be one I can perform given no input except the sentences themselves. In short, it seems that in doing this job I am implicitly relying on something like an effective

procedure."<sup>20</sup> In Chapter 4 I will expand this line of thought and produce further arguments to show that although the procedures which guide our use of language are not obvious, and need to be discovered, this does not count against their existence. I will argue that we must postulate such procedures (or conventions, at least) if we are to explain why there is widespread agreement on the use of a word with respect to an open class.

But regardless of whether my later arguments are successful, I think it is clearly possible (at least) that the <u>criterion</u> (in Putnam's sense), which people use when they apply certain words, is not always obvious, or trivial, and may be complex and interesting. For this reason I think it is sheer dogma on Putnam's part to claim that there are no hidden synonymies to be discovered. Clearly a lot more would need to be said about <u>criteria</u> to enable us to draw any firm conclusions on this question.

(It is interesting to note the similiarity between Putnam's use of 'criteria' and Bennett's claim that certain highly indispensible sentences provide the <u>test</u> for Fness in virtue of which other sentences about Fness come to be established. I think both of these notions are closely related to the idea that there are conventions which guide our use of language. It is this latter idea that I will pursue in the chapters which follow.)

#### Chapter 2

<sup>1</sup> Quine, From a Logical Point of View, p. 37.

- <sup>2</sup> Grice and Strawson, "In Defense of a Dogma", reprinted in <u>Necessity</u>, ed. Sumner and Woods, New York, 1969, p. 143.
- <sup>3</sup> Gilbert Harman, "Quine on Meaning and Existence", <u>The Review of Metaphysics</u>, vol. XXI, 1967, p. 137.
- <sup>4</sup> Quine, <u>Op. Cit.</u>, p. 32.
- <sup>5</sup> Grice and Strawson, Op. Cit., p. 152.

<sup>6</sup> Jonathan Bennett, "Analytic-Synthetic", reprinted in <u>Necessity</u>, p. 175.

<sup>7</sup> Quine, Op. Cit., p. 26.

- <sup>8</sup> Grice and Strawson, <u>Op. Cit.</u>, p. 153.
- <sup>9</sup> <u>Ibid</u>., p. 158.
- <sup>10</sup> Bennett, <u>Op. Cit</u>., p. 164.
- <sup>11</sup> Ibid., p. 165.
- <sup>12</sup> Ibid., pp. 165-6.
- <sup>13</sup> Ibid., p. 162.
- <sup>14</sup> Ibid., p. 163.
- 15 Ibid., p. 169.
- <sup>16</sup> Ibid., p. 177
- <sup>17</sup> Hilary Putnam, "The Analytic and the Synthetic", <u>Minnesota Studies in the Philosophy of Science</u>, III, ed. by H. Feigle and G. Maxwell, Minneapolis, 1962, p. 360.
- <sup>18</sup> Ibid., p. 362.
- 19 Ibid., pp. 392-93.
- 20 Putnam, "Some Issues in the Theory of Grammar", <u>Proceedings of the Twelfth</u> <u>Symposium in Applied Mathematics</u>, American Mathematical Society, Providence, 1961, p. 39.

Chapter 3

Conceptual Revision

"Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system. Even a statement very close to the periphery can be held true in the face of recalcitrant experience by pleading hallucination or by amending certain statements of the kind called logical laws. Conversely, by the same token, no statement is immune to revision."<sup>1</sup> In this passage, and in its adjacent paragraphs, Quine is claiming that in the face of an experience which conflicts with an accepted theory absolutely any sentence might be given up as false, provided sufficient changes were made elsewhere in our system of beliefs. What I wish to show in this chapter is that some sentences, at least, cannot be rejected unless we either make a mistake or change the meaning of the sentence. (i.e. change the assertion made by the sentence). My proof is as follows.

1. If a sentence S asserts the same thing, or can generally be used to express the same belief, at  $T_1$  as at  $T_2$ , then if S is true at  $T_1$ , then, a) S must be true at  $T_2$  and

b) the denial of S must be false at  $T_2$  and

c) whoever denies S at T2 is wrong.

2. If the sentence P which presently asserts the law of <u>Modus Ponens</u> is true, then whoever denies P, at any time, is either wrong or does not really deny what the sentence P presently asserts.

3. The law of <u>Modus Ponens</u> is true, and the sentence which presently asserts Modus Ponens is true.

4. Whoever denies, at any future time, the sentence which we presently use to assert <u>Modus Ponens</u> is either wrong or he does not deny what we presently assert.

On a certain usage of 'asserts the same thing', premise 1 admits of exceptions. For example, a token-reflective statement such as "I am happy" may have different truth values when asserted by different people, or when asserted by the same person at different times, even though, on a certain understanding of 'assert the same thing', I assert the same thing when I utter "I am happy" at different times.

However, on my use of 'asserts the same thing' premise 1 does not admit of these exceptions. There is a standard philosophical usage (which I am here following) according to which it is correct to say that when I assert "I am happy" and you assert "I am happy" we are asserting different things. On this usage a minimal condition for saying that two sentences make the same assertion (or express the same belief) is that they have the same truth value.

Regarding premise 3, if anyone chooses to deny premise 3 I do not wish to argue with that person, indeed I cannot argue with that person. The reader may wonder whether Quine's position is affected by the argument just given. I think it is. For I have produced a case where it is clear that we could not reject a particular sentence without either making a mistake or changing its truth value (and hence its meaning and assertive content). And in the latter case, where we reject S by changing its truth value and its assertive content, it is clear that we have not rejected the assertion made by S at  $T_1$ , since we still believe S was true at  $T_1$ . It follows that we cannot reject the assertion made by <u>some</u> sentences, e.g., S, without making a mistake. But this is just what Quine denies when he says, "No statement is immune to revision."

Quine might try to counter this argument by insisting that no real

distinction exists between (a) cases where rejection occurs because a sentence which we formerly thought to be true is discovered to have been false all along, and (b) cases where rejection occurs because the truth value of a sentence is being changed by changing the meaning of the sentence. But this counter will not work in the present case because our evidence for saying that the meaning of S has changed would be that its truth value had changed. One might claim that, in general, no real distinction exists between saying that our <u>opinion</u> about the truth value of S has changed and saying that the truth value of S has been changed. But this would entail the absurd position that no real distinction exists between saying that a sentence S is false at  $T_1$  and  $T_2$ , and saying that S is true at  $T_1$  and false at  $T_2$ . It would also entail the false position that no real distinction exists between what we believe to be the case and what is the case.

As I mentioned in the last chapter, Bennett has offered a Quinean account of how to distinguish cases where we should say that S is now false and has been false all along, though we once thought it to be true. He claims that if S was highly indispensible in our network of beliefs (i.e. analytic on Bennett's theory), then we should say that S was true but now is false. If S was not indispensible in our network of beliefs (i.e. not analytic on Bennett's theory), then we should say that S was false all along though we once thought it to be true.

For reasons which were given in the preceding chapter, I find Bennett's account incomplete. Consequently, I would like another account of the distinction between sentences which are "True then, false now" and sentences which are "False all along though we didn't know it". I think I can

provide a way of telling, <u>in some cases</u> at least, whether we should say that the possible future rejection of some sentence S, which we now believe would prove us wrong, or whether we should say that future people assert something different than we do. (a) If S is now used to assert something which is <u>obviously</u> true (e.g. there are people, people eat breakfast, 2+2=4, or -(p·-p) ), and future people regard S as <u>obviously</u> false, this would show that future people use S to assert something different from what we now assert. (b) If S is not now used to assert something obviously true, then we may suppose we were mistaken in our present belief in S.

Now someone might think that in appealing to the notion of <u>obvious</u> truth I am presupposing that some of our beliefs are infallible. But I am not. It is <u>possible</u> (perhaps) that we are mistaken in thinking that any given sentence is true. But this should not prevent us from saying that the truth value (and assertive content) of a sentence has changed, any more than it should prevent us from saying that <u>any</u> object has changed. Any belief of the form "This object has changed for x to y" is fallible. Nevertheless, it is obvious that things do change. There is just as strong a case for saying that the truth value of particular sentences change as there is for saying that anything changes.

The fact that it is possible that our most certain beliefs may be mistaken does not mean that these beliefs cannot be used in philosophical argument. Any of Quine's premises <u>could</u> be false. That does not mean that they are not well enough known to be used as a premise in an argument. The same holds true of premises like the law of <u>Modus Ponens</u> and "There are people". These might turn out to be false, but <u>in fact</u> they will not. This

is not dogmatism. It is merely the recognition that in philosophical discussion we are entitled to premises. And if we are entitled to any premises we are entitled to these. It would be methodologically absurd to suspend judgement about these and to continue an intellectual inquiry. This methodological absurdity is illustrated by the oddness of saying "It is logically possible (consistent with the laws of logic) that we are mistaken in thinking the laws of logic are true."

In my argument at the beginning of this chapter I say that whoever denies the sentence, S, which I now use to assert Modus Ponens is either wrong or does not deny what I have been asserting. Actually we can decide between these alternatives. It would be much more reasonable to suppose that one who denied Modus Ponens was not denying what we assert than it would be to suppose such a denial to be wrong. This is because it is much more likely that people should change meanings than that they should be mistaken about something as obviously true as Modus Ponens. This conclusion is very close to some of Quine's remarks where he says "Pre-logicality is a trait injected by bad translators."<sup>2</sup> Quine here implies that an assumption which should guide us in the translation of any hitherto unknown language is the assumption that speakers of the language do not hold beliefs involving elementary logical mistakes. This is not because it is impossible that any beings might make elementary logical or facutal mistakes, but rather that we always have better reason to suppose that our translation is wrong than that a linguistic community might make very elementary logical mistakes.

What I have said so far is compatible with a kind of Quinean position which may come close to the theory which Quine sketches at the end of

"Two Dogmas of Empiricism" and which is still acceptable to many traditional empiricists. (This position is suggested by Grice and Strawson). The modified Quinean position goes like this: "Our present set of concepts may not be totally adequate to cope with describing reality. Even our present logical concepts might be improved on, and perhaps even our concept of truth may be too coarse for describing reality. Our total set of concepts might be likened to a grid of squares which we use to approximate a curve. An improved set of concepts would be more closely grained: the grid of squares would be smaller and the curve more closely approximated. (For example, a language which was not of subject-predicate form, and which did not involve subject-predicate concepts might be a better tool for understanding reality than our present language.) It is hard to imagine a set of logical concepts which could enable us to better deal with reality than our present ones, but for all we know they could exist. And these concepts might just be linear descendents of our present concepts, and be expressed by the same words. Propositions might be expressed which are closely related to the propositions we now express. These closely related propositions might be expressed by the same sentences which now express their close relatives. For example, in the future people might use the words 'if p, then q' to express what we would express by 'if p, then probably q'. And the other truth functional connectives might undergo similar changes. In this language of the future an inference of the form "If p, then q. p, therefore q" would be invalid, since it would not follow that q was true - only that q was probably true. It might be that this probabilistic concept of 'if-then' and its related truth-functional concepts would enable us to better cope with reality

than our present concepts, and we might adopt these new concepts for this reason. We might reject our old concepts, and "reject" <u>Modus Ponens</u> in the sense that we no longer used <u>Modus Ponens</u>. This would not show that <u>Modus</u> <u>Ponens</u> expressed an invalid inference in our old language, or that people who once used <u>Modus Ponens</u> were making a mistake, but it would mean that a more powerful or more efficient system of concepts and inferences had been developed which was replacing our old system."

In the above I tried to sketch a way in which one could be said to reject or abandon our present concepts. But there is something peculiar This set of new, though related, concepts which are about this sketch. supposed to replace our present concepts can already be expressed in our present language (and this must be true of any example we can describe) and can perfectly happily coexist with our present set of concepts. Thus there seems to be no need for our present words to change meaning, even slightly, in order to express these new concepts. The question arises, therefore, how experience might influence us to change the meanings of our present words. Putnam, in "The Analytic and the Synthetic" answers this question with an example. His example is roughly as follows: Suppose that the word 'bachelor' now means "a sane male adult who has never been married". And suppose also that at some future time it is discovered that all bachelors have some neurosis, call it sexual frustration. (A very unlikely example - but that's another story). As far as anyone knows all and only bachelors suffer from sexual frustration. And imagine that everyone acquires such a high degree of psychological insight that they can tell within a few minutes conversation with a person whether he is sexually frustrated, and hence, whether he is a

bachelor. Also suppose that a whole cluster of interesting psychological laws are discovered about sexually frustrated people. Now under these circumstances some factual discovery might lead us to change the meaning of 'bachelor'. For example, if it were discovered that in rare cases there are bachelors who are not sexually frustrated we should either have to give up a whole cluster of completely general psychological laws about bachelors, or else change the extension (and meaning) of 'bachelor' slightly so that 'bachelor' referred to all and only sexually frustrated people. Which of these alternatives we choose will depend upon how interested we are in the old concept of a bachelor, and how difficult it would be to rephrase a whole cluster of psychological laws in terms of some new word. If we lose interest in the old concept of a bachelor (because marriage became an infrequent occurrence, for example) we might well change the extension of the word 'bachelor' to allow "All bachelors are sexually frustrated" and a whole cluster of psychological laws to remain exceptionless. In that case the old law, "All bachelors are unmarried" would come to have exceptions. However, this would not mean that people were formerly mistaken when they asserted that all bachelors are unmarried. For since the extension of the word 'bachelor' would have changed there is no reason for the old law to have the same truth value as the new law, Some may question whether we should have good reason for saying that the extension of the word 'bachelor' had changed. My answer to this is that we would have the best possible evidence for saying that the extension had changed, namely that "All bachelors are unmarried" had changed from a true law to a generality with exceptions.

# Chapter 3

<sup>1</sup> Quine, <u>From a Logical Point of View</u>, p. 43.

<sup>2</sup> Quine, <u>The Ways of Paradox</u>, New York, 1966, p. 102.

# Chapter 4

Concept Analysis and the A-S Distinction

.

In this chapter I will try to give a theoretical basis to the analytic-synthetic distinction. In order to do that I will first describe what I think is a very useful kind of concept analysis, and then go on to define the analytic-synthetic distinction in terms of this special kind of concept analysis. Before doing either of these, however, I would like to take a look at some familiar notions of concept analysis. It is often assumed by philosophers that there is a large amount of agreement about what counts as a concept analysis, and about what expressions are synonymous, but that we lack a theoretical basis for making these distinctions. I think the problem is more complicated than that. There is in fact a large amount of disagreement about specific cases, and that greatly complicates any attempt to give these distinctions a theoretical basis.

Suppose we temporarily put aside Quine's doubts about synonomy, necessity, and concept analysis. And suppose we agree that "2" and " $\sqrt[5]{32}$ " are necessarily coextensive concepts. Would we also say that " $\sqrt[5]{32}$ " is an analysis of our concept of "2"? And would we say that '2' and ' $\sqrt[5]{32}$ ' are synonymous expressions? Some philosophers would. Some philosophers are willing to count any description which has structure or complexity as an analysis if it is necessarily coextensive with the concept to be analysed. I do not wish to argue with these philosophers. Perhaps there are a sufficient number of philosophers to justify that use of 'concept analysis'. I do think, however, that there are finer and more useful distinctions to be drawn than that between necessary and non-necessary truths.

Some philosophers, on the other hand, reject the idea that  $\sqrt[4]{32}$  expresses an analysis of our concept of "2", or is synonymous with '2',

because, they say, an expression cannot mean anything more than what people usually mean by it. It seems implausible to hold that when one says that there are two birds in the tree, one means that there are the fifth root of 32 birds in the tree. In general, it is claimed, one cannot mean that x when asserting y, unless one believes y only if one believes x. On this account two expressions are synonymous only if they are interchangeable with one another in belief contexts, and the expression of the correct analysis of a concept must be interchangeable with the expression of the concept to be analysed in belief contexts.

The view just expressed comes close to a view discussed by Benson Mates in his article "Synonymity".<sup>1</sup> It would be convenient if this view were true, because the criterion of synonymity expressed is clear and simple. Unfortunately, there is no reason to believe that the proposed criterion is correct. This becomes apparent once we realize that we rarely know the analysis of concepts we normally use. (For example, most people could not produce an analysis of our concept of "grammatical" when asked, nor could they easily recognize its analysis when presented with it. In general, it is a difficult task to produce a correct concept analysis, and it often requires considerable reflection to ascertain whether one has arrived at the correct result.) Given this fact, there is no reason to think that the analysis of a concept will be interchangeable with the expression of the analysed concept in belief contexts.

It might be objected that in some sense of 'know' we unconsciously know the analysis of all the concepts we use. But this sense of know, if it exists at all (which I doubt), is not going to be helpful in the present

case. For in this sense of 'know' it is possible to know that p and yet be unaware that p is true. In this sense of know someone could "know" that p and yet sincerely assert "I do not believe that p." So the fact that someone "knows" the analysis of a concept does not give us any reason to think that the analysis will be interchangeable with the analysandum in belief contexts, unless we resort to talking about what a person unconsciously believes. If we move in this direction, however, the "substitutivity in belief contexts" criterion for synonomy becomes highly mysterious and problematical. I do not think the move from "synonomy" to "unconscious beliefs" is a move in the direction of clarity.

Church has argued in "Intensional Isomorphism and Identity of Belief"<sup>2</sup> that in fact the expression of a concept and its analysis are always interchangeable in belief contexts. His argument, briefly, is this: "There are many words in, say, English which can only be translated into another language, say, German, by analysing the English concept into its components, and then translating the analysis. For example, there is no single German word which means fortnight. Consequently, one must translate the word 'fortnight' into the German expression for "two weeks", and the German translation of "John believes a fortnight has passed" (E) is the same as the German translation of "John believes two weeks have passed" (E'). Call the German translation G. Church then claims that since any translation has the same truth value as the sentence being translated it follows that G must have the same truth value as both E and E'. Hence E and E' must have the same truth value in English. But E' is just the result of replacing the expression of a concept in E with its analysis, and in principle "fortnight" and "two weeks" could have been any

English expression and its concept analysis. Therefore, in principle, the same argument could be given to show that the expression of any concept and its analysis are interchangeable in belief contexts <u>salva veritate</u>."

Church's argument can quickly be shown to be inadequate. For Church gives no reason to think that the translation of the English belief context, E, into the German belief context, G, must have the same truth value. Presumably he would argue that E and G must have the same truth value because they mean the same. But there is no more reason for saying that E and G mean the same than there is for saying that E and E' mean the same. Consequently, anyone who thinks that E and E' have different truth values will remain unmoved by Church's argument, since such a person would either deny that E and E' mean exactly the same, or would deny that belief contexts which mean the same must have the same truth value. This would be more obvious if E were "Mary believes the theory is true" and if E' were "Mary believes the theory is  $\emptyset''$ . (where ' $\emptyset'$ ' is a description of the correct analysis of our concept of truth.) Actually, Church's argument is an unsuccessful attempt to solve the famous paradox of analysis. The paradox is usually stated somewhat It is thought that any analysis of a concept expresses the same as follows: concept as the concept being analysed. From this it follows (or is usually thought to follow) that any sentence, S, containing a particular expression, E, expresses the same proposition as the sentence which results from substituting for E in S the concept analysis of E. Thus, suppose that A is an analysis of the concept expressed by E. Then it ought to be true that 'E = E' expresses the same proposition as 'E = A'. But that seems absurd, since 'E = A' expresses a concept analysis whereas 'E = E' does not.

Furthermore, it is doubtful that whoever believes that E=E also believes that E=A. The problem, then, is to reconcile these facts with the claim that 'A' and 'E' express identical concepts.

I will now try to describe a kind of concept analysis which (1) can explain the paradox of analysis, (2) can explain why the term expressing a concept applies in just those cases it does apply, and (3) can explain to some extent how we recognize that a term, and the concept it expresses, apply in a given situation. The kind of concept analyses I will describe constitutes a subclass of those statements which would be classed as concept analyses by the account given in terms of the necessary coextensiveness of concepts, and will contain as a subclass those statements classed as analyses by the interchangeability in belief contexts criterion. This means that my account will be more restrictive than the "necessary coextensiveness" criterion, but not as restrictive as the "interchangability in belief contexts" criterion.

My account is given in terms of linguistic conventions or rules. I will argue that such conventions must exist if we are to explain why we apply language in the situations we do apply it, and if we are to explain the regularity that language exhibits. Of course, there is a very obvious sense in which the relationship between a word and its denotation is conventional. There is nothing intrinsic about the word 'rain' which makes us use it as we do. We could just as well have interchanged the roles of the words 'rain' and 'snow' in our language. In what follows I will try to connect this element of conventionality with the notion of a linguistic rule and with concept analysis.

Now some philosophers have suggested that concept analyses ought to state the rules which govern our use of words, and some philosophers (Kant,

Bennett) have said that concepts are rules, or sets of rules, for applying words or classifying objects, situations, etc. In the case of terms applying to physical objects these rules or conventions may be vague or imprecise, or incomplete, but in the case of logical terms, numbers, and other abstract notions such as philosophers are prone to deal with, the rules may be very precise and exactly stateable.

But some philosophers have ridiculed the idea that there unconscious rules governing our application of words to situations, or objects. Ziff, for example, claims that behavioural regularity is rule-guided only if the regularity is the result of conscious intention - only if the regularity is planned. Ziff scorns the idea that the regularity which language exhibits is rule-guided. At no time did people ever sit down and draw up rules for language, and at no time did people ever plan to achieve a certain end by inventing language.

I want to grant Ziff the point that linguistic regularity is not always, at least, the result of planning or conscious intention, but deny that this fact counts against saying that linguistic regularity is rule-guided. I maintain that behaviour does not have to be the result of conscious intention in order to be rule-guided. Perhaps rule-guided behaviour must be purposeful, but it is a mistake to think that all purposeful behaviour is consciously intended. Now certainly linguistic regularity exists for a purpose, namely, communication. Ziff tries to obscure this point where he ways, "The importance of communication is usually exaggerated",<sup>3</sup> and he goes on to produce examples of linguistic utterances which he alledges are not intended to communicate anything. But Ziff must admit, if we are to take him seriously, that most assertions are made for the purpose of communicating something, at least.

I will assume that most linguistic behaviour is purposeful. Given this premise I think I can build a good case for saying that language is rule-guided. I base this claim on the fact that language does satisfy an analysis of rule-guidedness which can easily be extrapolated from David Lewis' analysis of rules in <u>Convention</u>,<sup>4</sup> and I think Lewis' analysis is correct. Lewis' account of rule-guidedness is roughly as follows: A behaviour pattern  $P_1$  is rule-guided if and only if (a)  $P_1$  is a voluntary system of regularities which are performed by members of a group in order to achieve some end (e.g. communication) which is mutually desired by all members of that group, (b) that same end (e.g. communication) could be achieved if members of the group had each chosen to act according to some different behaviour pattern,  $P_2$ , and (c) it matters little to members of the group whether they act according to  $P_1$  or  $P_2$ , but it matters to each member of the group that he/she shall act according to whichever of the two patterns most other members adopt.

Condition (c) insures the kind of arbitrariness or conventionality required for rule-guidedness. The idea is that members of the group are largely indifferent to what behavioural pattern they follow as long as the desired result occurs. It matters little to any of us, for example, whether we call something 'red' or 'sned' as long as we succeed in communication. Conditions a, b, and c are all met by human languages. Consequently, there is no barrier to saying that our linguistic behaviour is rule-guided (assuming Lewis' analysis is right).

We are now in a position to sketch what I think is the interesting and perhaps most useful conception of a concept analysis. A correct concept

analysis states a set of rules for applying words to objects or situations (or for distinguishing objects and situations in the non-linguistic case) (a) which will select all and only those objects or situations which satisfy the concept being analysed, (b) on the basis of which we actually do apply the expressions which express the concept being analysed, i.e., the rules actually do guide our use of these expressions in something like the way a computer's response to external stimuli is guided by its program.

(In order to be as clear as possible, I would like to explain the analogy between rule-guided behaviour and computer-programmed behaviour. It is important that we have some explanation of how rules guide our linguistic behaviour, otherwise it will not explain much to say that our linguistic behaviour is rule-guided. So here is my explanation. On my theory, linguistic rules constitute a partial program of our brains. This programming could occur if each brain were innately programmed to program itself further, according to the particular linguistic environment which it grows up in. There is nothing absurd in this idea. A computer could be programmed to further program itself according to the kind of environment it was placed in, and there is no reason why humans could not do the same thing, at a neural level. It seems plausible that children do something like this when they learn the grammar of the language they learn to speak.

Certain neural changes may occur when we learn to use a particular word, and these neural changes may cause us to use the word as if we were consciously following a certain rule. In such a case it may be appropriate to say that the rule is guiding our use of the word. For example, if (a) a creature's neural structure changes as a result of the creature's having

observed some linguistic regularity which holds by convention (in Lewis' sense), and if (b) the linguistic behaviour produced by that neural change is such as would be produced by consciously following a particular rule, then we may say that the linguistic behaviour is guided by the linguistic rule. We may say this because (1) the relevant linguistic behaviour is being guided by a neural change which occurred in order to enable the creature to imitate a linguistic regularity, and (2) the linguistic regularity being imitated is consequently determining the relevant linguistic behaviour, and (3) the linguistic regularity being imitated is a regularity which exists by convention (in Lewis' sense). This, I suggest, is the mechanism by which conventions guide linguistic behaviour.)

More has to be said about my sketch of concept analyses but first let us see how it satisfies the requirements I laid down. (1) The concept analysis explains why the term expressing a concept applies in just those cases it does apply, because the concept analysis states the rules for applying the term and the rules link the term to just those cases to which the term can be correctly applied. (2) The concept analysis explains to some extent how we recognize that a term applies in a given situation, because the analysis states the rules which guide our linguistic response in that situation. We come to know whether the term applies to a given case because our response to the case is guided by rules which we have learned in language training. (3) My account of a concept analysis explains the paradox of analysis. To see this consider the following.

On my account a concept analysis describes the rules which guide the use of some word. The rules which guide our use of words need not be conscious

rules that we could formulate on demand; they need not be applied with our awareness. We can know how to use words without knowing what we are doing when we use them, just as a musician may know how to play a piece without being able to describe the complex ways she is moving her fingers. This explains how we can know how to use a concept without knowing the analysis of that concept.

The rule-guidedness theory of linguistic behaviour can also explain how a person can know the sentence "This sentence is grammatical" to be true without knowing "This sentence is XYZ" to be true (where 'XYZ' is some complex description which is necessarily coextensive with the word 'grammatical'). The explanation is this: The denotation of the complex expression 'XYZ' is a function of the rules guiding our use of 'X', our use of 'Y', and our use of 'Z' (assuming 'X', 'Y', and 'Z' to be separate words or phrases). These rules are different from the rule guiding our use of 'grammatical', even though the combined effect of these rules logically insures that the expression 'XYZ' applies to just those sentences that 'grammatical' applies to. And the fact that the rules guiding 'XYZ' are different from the rule guiding 'grammatical' allows a person to correctly apply the sentence "This sentence is grammatical" in a given situation without knowing whether the sentence "This sentence is XYZ" would apply in that situation. In general the rules guiding the use of a simple expression are different from the rules guiding the use of a complex expression. That is why the expression of a concept and the analysis of that concept are not interchangeable in belief contexts. Once we see the concept expressed by a word as a rule we are no longer tempted to think that the analysans and analysandum phrases express the same concept.

We can also see why it is misleading to say that simple sentences and complex sentences express the same thought or proposition. And insofar as it makes sense to say that two words in different languages express the same concept it is because it makes sense to say that speakers of different languages are following practically the same linguistic rules. The words are different, but the rules are the same, in the relevant way.

I think I have now explained the paradox of analysis. Much more could be said about it, but this is not the place.

## (aside)

It is of interest to note that sometimes we can recognize a concept analysis as correct when it is presented to us. Something analogous happens with nonlinguistic behaviour. It sometimes happens that we recognize that the reason we acted in a particular way is that we were following a certain rule. For example, we might ask a tennis player whether the reason she acted in a particular way in a given situation was that she had learned a rule to the effect that in situations of such-and-such a kind it is best to act so-and-so. Our tennis player might correctly answer yes, even though she could not have produced the rule which would explain her action.

### (end aside)

I turn now to some problems with my sketch of a correct concept analysis. An immediate problem arises when we try to decide which of the logically equivalent alternative analyses actually states the true analysis of a concept. That is, which of the alternative proposed analyses actually states the rules which guide our application of words to objects or situations? If these rules are unconscious how are we to decide what rules guide our

linguistic behaviour? For example, our use of the expression 'grammatical in English' seems to be rule-guided. There is widespread agreement about which sentences are grammatical and which are not. But if we ask ourselves whether a particular transformational grammar in fact states all the rules which govern our use of the word 'grammatical' we draw a blank. Parts of the grammar may be easily recognised as correct, e.g., a rule which says that possessives such as 'the brother of John' may be converted into 'John's brother', but this is not true of more complex rules pertaining to the deep structure of complex sentences. The question arises, how could we decide which (if any) of two complete transformational grammars states the rules which actually guide our judgements of grammaticality? Perhaps this question can be answered as follows.

It seems in principle possible that we should at some future time be able to trace the circuitry of the brain so completely that we could obtain a map or flow chart of the successive electronic states of the brain when a person generates a given sentence, or when a person performs a given calculation. And it seems possible that there could be an isomophic mapping between the successive transformational stages which a grammar assigns to the generation of a particular sentence and the successive brain states which occur when that sentence is generated.

To see that this is possible consider an analogous case with computers. It is well known that there is an isomorphic mapping between statements in the propositional calculus and circuits in computers. The formula 'p v q', for example, corresponds to a circuit which has two branches, like this -. If either of the two branches has a current running through it, i.e., is a closed circuit, then the major trunk has a current running through it. Analogously, if one of the pair p,q is true, then the whole disjunction is

Truth in a sentential component of a formula corresponds to a closed true. circuit in circuitry, falsity corresponds to an open circuit, and disjunction corresponds to a branched circuit. In general, given any formula in the propositional calculus one can draw a circuit corresponding to it, and viceversa. Consequently, since different derivations of the same formula in the sentential calculus correspond to different conditionals having different antecedents and the same consequent, it follows that different derivations of the same formula in the propositional calculus correspond to different circuits in computers. Analogously, if two computers were programmed to test sentences for grammaticality according to two non-isomorphic, but logically equivalent, grammars, then the processes which take place in each computer when the computers test a sentence would be different. And if we knew what processes were taking place in a computer when it tests a sentence for grammaticality we could tell how the computer was programmed. Likewise, I think, if we know what processes take place in the human brain when a person judges a sentence for grammaticality, we should be able to reconstruct a program for the brain which is isomorphic with the processes that take place in the brain. And if we find that a particular transformational grammar, G, is isomorphic (at a certain level of detail) with the program we have reconstructed for the brain, and if the individual sentences of the grammar G are isomorphic with the individual sentences of the reconstructed program with which they are paired, then we could conclude that G correctly represents the grammar which guides our judgements of the grammaticality of sentences. In this way we could decide which of two non-isomorphic, but logically equivalent grammars, if either, expressed the rules which constitute our concept of grammaticality.

Whether in fact we will ever be able to decide, in just the way I have described, which of two non-isomorphic, but logically equivalent grammars express our concept of grammaticality depends in part upon whether the brain is very much like a computer. But even if the brain is <u>not</u> very much like a computer, we might still decide this question is some way which is roughly analogous to the method I have described.

(It might be helpful at this point to stress that the example just produced appeals to the premise that we can infer the program of a computer from a description of its structure and its internal processes. The converse of this premise does not hold. From the fact that a computer is programmed a certain way we cannot infer that the computer has a particular internal structure. Two computers may follow the same program although they have differing internal structures. By analogy, the fact that two people are following the same rule does not allow us to infer that they have the same brain structure and brain processes. However, if two people <u>do</u> have the same brain structure and brain processes this constitutes evidence that they are following the same linguistic rules.

Also, if two computers, which have identical internal structures, go through different processes when solving the same problem, we may infer that they have been programmed differently. Analogously, if two people with identical brain structures (in the relevant respects) need to go through different intellectual processes when judging, say, the grammaticality of a sentence, we may infer that those two people are being guided by different rules.)

A problem arises now, and that is, isn't it implausible to suppose

that the correct analysis of our concept "grammatical in English" must await future psychological inquiry, and in general isn't it odd to suppose that concept analysis is an empirical enquiry which can yield only probablistic results? Perhaps so, but I think good reasons have been given to support this conclusion, and perhaps philosophers must, like other thinkers, be prepared to accept some surprising results. Arguments have been given to show that there have to be rules or conventions guiding our use of language, and it just is an empirical question what these rules are. Arguments have also been given to show that it explains a lot to suppose that concept analyses state the rules which guide the use of words, (or guide us when we make distinctions in the non-linguistic case). But I do not really insist that we use the expression 'concept analysis' in this way. I do not care how we use the expression 'concept analysis'. My main concern has been to show that we can draw a distinction which is finer and more useful that the necessary/non-necessary distinction, and yet which is not as restrictive as the interchangeability in belief contexts criterion for synonomy. I think some definite epistemological advantages result from focusing our attention on the rules or conventions guiding the use of words. In addition to the merits already considered under points 1-3, there is the following point. By focusing our attention on rules governing the application of words, we can explain why all the objects or cases described by an expression have certain properties in common. We can explain why all the items in class C have the property P, if we can show that a rule for placing an item in class C is "the item must have property P". For example, it genuinely explains why all the objects in a certain box have four sides if we know that a rule for

placing an object in that box is "the object must have four sides". Further explanation is required to explain why all the objects in the box have the cube root of 64 sides.

To explain this latter fact we must talk about entailments between arithmetical propositions. So there is an advantage to focusing our attention on the rules governing the use of language that does not result from merely focusing our attention on the necessary coextension of concepts or the necessary entailments between propositions. To analyse a concept into properties P, Q, and R does not explain why all the items satisfying the concept have properties P, Q, and R unless we point out that the concept is a rule for selecting items according to whether they have the properties P, Q, and R.

An objection to my last series of points may be made on the following grounds. What is to prevent us from saying that a rule to the effect, say, that a word should be applied only to objects having four sides is the same as a rule to the effect that the same word should be applied only to objects having the cube root of 64 number of sides? If we can provide no reason for saying these are different rules, must we not conclude that my account of concept analysis has no advantages over the account in terms of necessary coextensiveness?

To reply to this objection: (1) No, even if we can provide no reason for saying these are different rules, it is still an advantage to see concepts as rules since it still explains the fact that all the things classed under a single term have the properties they share in common. (2) I think we can provide a reason for saying these are different rules, namely, one

would go through different processes in following these two rules. I submit the following as a general principle for individuating rules. "If <u>any</u> person would need to go through different behavioural, or intellectual (thought) processes when consciously following rule x and rule y, then x and y are different rules." For example, most people would go through different intellectual processes when following the rules (a) "Put four-sided objects in this box" and (b) "Put objects in the box having the cube root of 64 sides". When following rule (a) one might count the sides of objects or just glance at objects to see whether they were four-sided. More than this would be involved when an average person followed rule (b). The average person would have to calculate a couple of seconds when following rule (b) and then, after calculating, he or she would go through the same procedure one would go through in following rule (a).

In this example it is fairly easy to say that different intellectual processes would occur when an average person followed rules (a) and (b). Not every case will be so easy to decide. There may be some vagueness in the notion of "same process". And for some purposes we may want to count slightly different processes as the same process. But I think that the notion of "same process" is sufficiently clear to enable us, in many cases, to distinguish logically equivalent rules. That is sufficient for present purposes. It is also interesting to note that we do not always need to be in a position to say what rule a person is following in order to have strong reasons for saying that the person is following some rule (just as we do not always need to be in a position to say <u>what</u> program a computer is following in order to know that it is following some program).

I turn now to a question which may have occurred to the reader; it is this: "How can we tell whether people who speak the same language are following the same linguistic rules?" Before attempting to answer this question we should note that we occasionally have strong evidence that people are following <u>different</u> linguistic rules. Practically everyone has had the experience of arguing with someone and discovering that the disagreement was <u>really</u> a verbal disagreement - that each party to the disagreement had a different use for some word. Often these verbal disagreements are uncovered by comparing metalinguistic remarks such as "Well, what would you count as a whale?" In examples such as these we can usually tell whether two people have extensionally different rules.

Sometimes, however, people make metalinguistic remarks about the <u>meanings</u> of words and about what sentences they take to be <u>analytic</u>. The fact that these intensional remarks often agree is easily explained by the <u>hypo-thesis</u> that many people follow the same linguistic rules. But it is very difficult (at least) to explain this fact by appeal to extensional considerations alone. The fact that the rule-guidedness hypothesis easily explains this uniformity of intensional metalinguistic remarks, combined with the lack of competing explanations, gives us some reason to accept this hypothesis.

People will usually complete the number series "2,4,6,8..." in the same way, even though there are infinitely many rules which could generate this sequence of numbers. This fact constitutes evidence that people's minds (or brains) often work in a similiar fashion. Also there are a small number of logical systems which people find easy and natural to work with, although

there are infinitely many ways to construct, say, a system of the predicate calculus. This further supports the view that most people have similiar intellectual equipment - that their "computers" are very similiar. Furthermore, what physiological evidence we presently have indicates that people's brains and brain processes are similiar. These facts suggest that people often are being guided by the same linguistic rules when their linguistic behaviour is identical in the relevant respects. For if people have structurally identical intellectual equipment, and their behaviour is identical, then they are probably undergoing the same internal processes. In principle, we could verify whether people do <u>in fact</u> undergo similiar brain processes when they judge, say, the grammaticality of the same sentence. If we found the brain processes, brain structure, and linguistic behaviour to be identical in the relevant respects we could say for sure whether people are being guided by the same rules.

We are nearly in a position to give a definition of the analyticsynthetic distinction. But first I wish to introduce a technical term, namely, 'the semantic transformation of a semantic rule', which may be understood somewhat as follows: In general the semantic rules which are described in a concept analysis will relate an expression, E, to a class of situations or cases, C. The semantic transformations of a given rule is the universal closure of a conditional, in which the antecedent asserts that situation C obtains, and the consequent expresses C in terms of the expression

E, according to the semantical rule in question. The following examples may be helpful:

1. Rule, "You may apply 'sister' to any female sibling."

Semantic Transformation of Rule<sub>1</sub>, "For any x, if x is a female sibling, then x is a sister."

2. Rule<sub>2</sub>, "You may connect any pair of true sentences with the word 'and'."

```
Semantic Transformation, "For all p and q, if p is true and q is true, then p and q is true."
```

Although this characterization is somewhat vague, it should give the reader a general idea of how to form the semantic transformation (S-T) of a given rule. The motivation for introducing the idea of an S-T is this: by considering the logical consequences of a given S-T we can uncover the purely logical consequences of following a given linguistic rule. This brings us to my definition of an analytic sentence, namely, "A sentence is analytic if and only if it is a logical consequence of the semantic transformation of the rules described in some concept analysis." "If a sentence is not analytic is it synthetic".

I think that my definition of the analytic-synthetic distinction nearly captures one standard conception of the distinction. I do not claim that it accurately matches everyone's use of 'analytic'. In particular it does not match that use of 'analytic' according to which 'analytic' is trivially interchangeable with 'necessary'. My definition of 'analytic' is designed to caputre the idea that analytic sentences relate to the analysis of a concept, or are in some way the logical consequence of the rules of language. The question whether, on my use of 'analytic', all necessary truths are analytic is an interesting question which I will consider in a later
chapter. (In what follows I will abbreviate my definition of 'analytic' by saying that a sentence is analytic iff it is a logical consequence of some concept analysis).

There is another question which I would like to consider at this stage, i.e., Does my definition make the analytic-synthetic distinction one of degree or of kind? Answer: In the sense that practically every distinction is one of degree the analytic-synthetic distinction is one of degree. But insofar as the question whether a sentence is analytic clearly depends on another question, e.g., what rules guide the use of this expression?, the distinction is a sharp one; it is one of kind. The questions of degree come in when we try to decide what the particular rules guiding the use of an expression are. Sometimes the rule guiding the use of an expression is vague, e.g., "Apply the term F to anything that has nearly all of the properties  $P_1...P_n$ ." This vagueness will engender undecidable questions about whether something is an F, but that should not lead us to suspect the concept of a rule, or any concepts involving the concept of a rule. Again, sometimes the question whether something is a rule guiding the use of a particular expression will be very difficult to decide, but that should not lead us to reject my analysis of analyticity in terms of rules. Likewise, questions about the truth of a sentence are sometimes undecidable, but that should not lead us to reject an analysis of knowledge in terms of true belief. Questions about truth are often unanswerable just because the use of some term is unclear, or because it is not clear what rules do guide the use of a term. For this reason I think Quine is wrong when he says that the concept of truth is respectable and tolerably clear, but that the concepts of meaning

and linguistic convention are neither respectable nor clear. Questions about truth can be settled only if there are definite conventions and rules guiding the use of terms. For insofar as the use of an expression is indeterminate, the use of any sentence containing that expression is indeterminate, and so is its truth value. And if there are definite conventions guiding the use of these terms, then questions about what these conventions are must be answerable in principle. I do not know of a decision procedure by which we can distinguish linguistic conventions from non-conventions, but then I have never seen a decision procedure by which we can distinguish true sentences from other sentences. I suggest, however, that we could never produce a decision procedure for truth unless we could produce a decision procedure for linguistic conventions.

#### Chapter 4

- Benson Mates, "Synonymity", <u>University of California Publications in</u> <u>Philosophy</u>, 1950.
- <sup>2</sup> Alonzo Church, "Intensional Isomorphism and the Identity of Belief", <u>Philosophical Studies</u>, 1954.
- <sup>3</sup> Ziff, <u>Semantic Analysis</u>, Ithica, 1960, p. 36.
- <sup>4</sup> Lewis, Convention.

# Chapter 5

· · ·

### Recalcitrant Experience

In chapter 3 I considered and rejected Quine's claim that absolutely any belief might be rejected in the face of an experiential report which conflicted with an accepted theory. In what immediately follows I hope to show that a counterexample to Quine's view is provided by one of his other views about theories.

In "Two Dogmas of Empiricism", and elsewhere, Quine expresses the belief that all theories must pass empirical tests, or be answerable to experience. This belief entails that there is a certain class of theories which adequately explains our experience, and another class which fails to adequately explain our experience. Quine also asserts that theories can and must be altered in the face of recalcitrant experience, and rightly so. For if we do not hold that some theories, unmodified, must be abandoned in the face of experience, then we can make no sense of the notion that theories must be answerable to experience, and so Quine's position becomes unintelligible. But if we admit that some theories, unless modified, must be abandoned in the face of experience by any rational being, then we admit that a certain belief is forced upon us by experience.

It is true, of course, that the <u>sentence</u> (S) "Some theories, unless modified, must be abandoned in the face of experience by any rational being", might at some future time be regarded as false by most people, but this would not show that this sentence is not now true. And it would be absurd for someone to reject the belief now expressed by S in order to save a theory. In much the same way it would be absurd to reject the law of non-contradiction in order to make one's theory consistent with experience.

Now an argument could easily be constructed (exactly analogous to

to the argument given to defend Modus Ponens) which would show that whoever denies S in the future is either wrong or does not deny the belief we now assert by S. We simply point out that assent to S is a precondition of intelligible discourse at this time. This enables us to take S as a true premise in any argument, and this commits us to saying that anyone who denies S in the future is either mistaken or else is not really denying what we now assert by S. The latter alternative might occur if new experiences lead us to modify our present set of concepts in some way, e.g., our concept of a theory, or our concept of experience, so that S comes to express something different from what S now expresses. What entitles us to say with assurance that S would express something different from what S now expresses is the fact that sameness of truth value is a precondition of sameness of belief or sameness of assertion. If S is false at some future time and S is true now, then S expresses something different in the future from what S expresses now.

So far we have shown that certain theories must be excluded by experience. But the problem still remains, how does experience disconfirm a particular theory in a particular formulation? What are recalcitrant experiences? Karl Popper has discussed this problem in Chapter V of <u>The Logic of</u> <u>Scientific Discovery</u>.<sup>1</sup> Popper introduces the problem by discussing a trilemma which is due to J.F. Fries. The trilemma is this. Either our beliefs about experience are to be accepted dogmatically, with no justification, or they are to be justified. If we accept the principle that any belief or statement can be justified only by deducing it from other beliefs or statements, then we cannot justify any statement without becoming involved in an infinite regress of justifications. But if we reject the principle

that all beliefs are to be justified by deducing them from other beliefs, then we must justify our beliefs some other way. Popper believes that the only other way beliefs could be justified is by experience itself (as opposed to a linguistic description of experience). Popper calls the doctrine that our beliefs are justified by experiences psychologism. Now although psychologism avoids the problems of both dogmatism and the infinite regress, Popper rejects psychologism because he thinks that experiences could only cause beliefs; not justify them. He believes that justification is a logical relation which can hold only between sentences or beliefs, and not between beliefs and experiences. In Popper's view causal explanations are never logical justifications.

Since Popper rejects psychologism he can only fall back either on dogmatism or the infinite regress. In fact he tries to combine these two alternatives. Briefly his theory of justification is as follows: In principle statements can only be justified in terms of other statements, and the possible chain of justification for a statement is infinite. But in practice we must stop somewhere in this chain of justification and just decide to assume that particular statements are true. Statements of this kind form the touchstone for science. They are the basic statements which describe intersubjectively observable events, and they must have a logical form such that their negations are inconsistent with general law-like statements. "Basic statements are therefore - in the material mode of speech - statements asserting that an observable event is occuring in a certain region of space and time."<sup>2</sup> In practice, we decide whether to accept or reject a particular basic statement after performing a test or experiment. We do not justify our decision to accept or reject a particular basic statement, however, since this would

involve us in an infinite regress. Rather we just decide, freely; though our decision is <u>rule-guided</u> and not arbitrary. Also our decision is not absolute. It can be changed in the light of future experience if we discover that some truth condition of a bæsic statement is not fulfilled. But once we have decided to accept or reject a basic statement (for the time being) we can use it to test our theories (for the time being). Thus our decisions to accept or reject basic statements justify our other beliefs, but these decisions and basic statements are not themselves justified.

That roughly is Popper's position. But Popper's view of basic statements is open to a serious objection. Popper says, "Basic statements are accepted as the result of a decision or agreement; and to that extent they are conventions. The decisions are reached in accordance with a procedure governed by rules."<sup>3</sup> (To that extent the decisions are not arbitrary conventions.) Now what is Popper to say if a group of scientists decide to accept a basic statement not on the basis of an observation, but on the basis of superstition, or religion, say? Popper will try to exclude this kind of case on the grounds that it violates the rules for accepting basic statements. One of these rules, presumably, will say that basic statements must be accepted on the basis of observation. But can Popper exclude basic statements whose acceptance violates certain rules, and also hold that our decision to accept basic statements need not be justified in any way? I think not. If a group of scientists adopted certain basic statements arbitrarily we would not accept their decision to adopt those statements. We would claim they were being unreasonable, and I think Popper would make this claim also, or some equivalent claim. To say that a decision to adopt a certain basic statement violated the

rules for adopting basic statements is to say that the adoption of that basic statement was unjustified.

Popper's solution is an attempt to deal with the problem of fallibility and revisability of all statements of objective fact. I think it is commonly recognised that all statements about the way things are (not the way things seem to be) are fallible. Sense-datum theorists and others have attempted to provide an incorrigible foundation for our knowledge of the external world. This foundation is supposed to consist of reports of experiences or sense-data. (Even these kinds of reports are fallible, however, since all reports about sensations involve comparison, and one can make a mistake in comparing any two things, even sensations.) Now even if one grants that sensation reports are infallible we must allow (and the sense-datum theorists do) that when we move from sensation reports to objective claims about the external world we make assumptions, inductive and otherwise, and these assumptions may just be false. I see no way to avoid admitting that beliefs about the way things objectively are fallible, but I do not see why this is especially a problem. Things do happen in the world, and partly as a consequence of this we adopt certain beliefs. These beliefs are fallible, of course, because sometimes beliefs about events are caused in us by things unconnected with the events we report. But we do have our beliefs, nevertheless, and when we believe that certain things have occurred which falsify some predictions of our theory, then we must modify our theory (provided we do not abandon our belief that the recalcitrant experiences occurred.) All that is required to make sense of the idea of recalcitrant experience is that we sometimes, collectively, have more confidence that a certain event E, or

class of events E, has occurred than we have that our theory is correct. Sometimes we are in this position. The question arises, how do we get in this position. The answer might be something like the following: 1. Suppose 2 people report having had E-like experiences (experiences such as they would have if E occurred and they were situated in such a way as to be causally affected by E).

Assume that when a person sincerely reports an E-like experience it is probable that an event of type E has occurred, say the probabilities are 10/1.
If 2 people report an E-like experience the odds are 99/100 that an E-type event occurred.

4. If 100 people report an E-like experience the odds approach certainty that E occurred.

5. Almost no theory has a probability approaching certainty.

6. In the case described (involving 100 people) we would have greater confindence that an E-type event occurred than that our theory is right.

The case just described is, of course, an oversimplification. Yet, I think, something like this kind of reasoning goes on when we test and reject theories in the light of our beliefs about experience. The procedure just described assumes the truth of certain of our experiential reports in order to establish the probable truth of our experiential reports. Does this raise infinite regress type problems? I think not. It would raise infinite regress problems if we first had to establish the truth of certain experiential reports in order to establish the probable truth of other reports, but I do not think we must do that. Rather we assume the truth of some of our experiential

beliefs on something like Kantian grounds, that is, on the grounds that if our experience is to be intelligible at all, then some of our experiential beliefs must be true.

(For purposes of this thesis I do not think it is necessary to defend the assumption that some of our experiential beliefs are true. But for curiosity's sake I will briefly outline a Kantian defense of this assumption. "If we did not assume that some of our experiential beliefs are true then we should have no beliefs about what constitutes experiences of an outer objective world, and no way of distinguishing our own subjective states from states of the external world. Hence, we should have no way of distinguishing ourselves from the outer objective world, and we would have no concept of ourselves. But since we do distinguish our inner states from an outer objective world, it follows that we must assume the truth of some experiential beliefs. Nothing dictates which of our experiential beliefs we must assume to be true, but we naturally retain those beliefs which enable us to form a coherent picture of the world.")

It would, of course, be circular to assume the truth of certain experiential beliefs in order to justify all our beliefs about experience, but I am not trying to do that. In fact I am denying that that can be done. The justification of all of our beliefs about experience is not something <u>we</u> can give. That does not imply that the justification of our beliefs about experience does not exist. The justification does exist, but it does not consist of a chain of reasons. In what follows I will try to give a rough explanation of what justifies our beliefs about the world. But this explanation should not itself be construed as a justification.

In an earlier chapter I argued that there are rules guiding our use of language. These rules guide our use of particular words, phrases,

and sentences. There are grammatical rules as well as semantical rules, and we often use or follow these rules without knowing that we do so. Sometimes these rules are uncovered in stating concept analyses. There are also rules tying language to experiences which can only be known ostensively. For example, there is a rule which may be roughly stated as follows: When you see something like <u>this</u> you may say "That is red." (Where the denotation of 'this' is given ostensively.)

The rules which guide our use of language do not <u>require us</u> to respond to our experiences linguistically. Rather they <u>allow us</u> to respond to our experiences with a certain range of descriptions. If there is no rule allowing a certain linguistic response, then that response is excluded by a closure rule which says that a response is permitted only if it is permitted by a rule.

We are now in a position to explain how linguistic descriptions of experience are justified. What justifies us in asserting a certain description of our experience is <u>the fact</u> (not the belief that it is a fact) that our linguistic response was a response to an experience and was guided by the rules governing the use of language, i.e., our response was linguistically acceptable.

The fact that our linguistic responses are rule-guided does not imply that we make decisions or form beliefs that we are following linguistic rules correctly, and it does not imply that we check to see that any conditions specified by the rules are satisfied. To follow a linguistic rule correctly it suffices to react to an experience in a certain way. This reacting to an experience may be a causal matter - the experience may be one causal condition

among other conditions which jointly result in a linguistic response. Other conditions (not necessarily causal) which must accompany an experience, E, in order for a correct linguistic response to E to occur may include such factors as willingness to respond linguistically, and previous language training, etc.

There might seem to be a problem in saying that a linguistic response to an experience, can both be partially caused (in the sense of 'provoked by'), and also rule-guided. But I don't think the problem is serious in this case. Suppose a computer is programmed to respond to stimuli according to certain rules when certain other causal factors are present. The programming of the computer according to certain rules then becomes one of the causal factors which jointly result in a particular linguistic response on the part of the computer. Thus the response of the computer is both provoked by external stimuli and guided by certain rules. Human beings, unlike computers, may not be programmed with all the rules they need to speak a language, but they learn these rules, and can be criticised if they violate these rules, even though their response may be provoked.

I think there is no more problem in supposing that human linguistic response may be both rule-guided and provoked by our experiences, than there is in supposing that the response of computers may both be programmed and provoked by external stimuli. In fact the case is easy to make for human beings, since a condition usually necessary for human linguistic response is willingness to respond linguistically. This means that linguistic response is usually voluntary. Thus, there can be no objection - to saying that linguistic responses are rule-guided - on the grounds that linguistic responses

are involuntary (though someone might object to calling computer behaviour rule-guided). The remaining conditions for saying that our linguistic behaviour is rule-guided are also satisfied (as I pointed out when I discussed David Lewis's account of rules in an earlier chapter. I repeat them now to refresh our memory. "A behaviour pattern,  $P_1$ , is rule-guided if (a)  $P_1$  is a voluntary system of regularities which are performed by members of a group in order to achieve some end (e.g. communication) which is mutually desired by all members of that group, (b) that same end (communication)could be achieved if members of the group had collectively chosen to act according to some different behavioural pattern,  $P_2$ , and (c) it matters little to members of the group whether they act according to  $P_1$  or  $P_2$ , but it matters to each member of the group that he/she shall act according to whichever of the two patterns most other members adopt.)

Now it may sound strange to say that linguistic responses to experiences are both provoked and also voluntary, and since rule-guided acts are voluntary one might think it strange to say that linguistic responses are both provoked and rule-guided. But there is no real problem here. The strangeness of these claims arises, I think, from thinking that whatever we call'responses' must be entirely caused by external stimuli, or must be correlatable in a 1-1 fashion with external stimuli. But I do not intend to use 'response' in such a narrow sense. When I say that some linguistic behaviour is a response to an experience, or is provoked by an experience, I <u>only</u> mean that the linguistic response will occur whenever the experience occurs, <u>provided</u> certain other conditions are satisfied. These other conditions include things like willingness to respond, having a reason to respond, having

language training, etc. I do not claim that these other conditions are partial <u>causes</u> of the linguistic response. That question is left open. Given this account of the sense in which linguistic responses are provoked by experience I see no incompatibility in saying both that linguistic responses are provoked and voluntary. If the reader chooses he or she may suppose I am using the expression 'provoked' in a technical sense. I neither intend to, nor am compelled to, take a stand on the question whether all human action is caused. (However, I am committed to saying that there is some set of internal and external conditions which will regularly be followed by a linguistic response. This seems a fairly safe claim once it is realized the internal conditions to which I refer include things like willingness to respond, having a reason to respond, etc.)

So far I have tried to explain the justification of some linguistic responses to our experiences or sense impressions. The question arises whether this kind of explanation can also be given to explain the justification of descriptions of objective states of affairs, e.g., "This is water". For descriptions of objective states of affairs go beyond what is justified by some descriptions of experiences. We all know that we may have experiences which are like those of water, without our experiences actually being of water. So what justifies our assertion "This is water"?

I think much the same kind of thing justifies descriptions of objective states as justifies descriptions of subjective states. Our assertion that a particular objective state exists is justified by the fact (not the belief that it is a fact) that our linguistic response was permitted by the rules governing the application of words. (Again, our reacting to an

experience in a rule-guided way need not involve any judgements on our part that the conditions for applying a rule are satisfied. Rather our response to an experience may simply be provoked by the experience.)

The rules which permit objective state-descriptions as responses to experiences are more complicated than those permitting subjective statedescriptions. Rules permitting objective state-descriptions will link linguistic responses to a whole network of experiences. For example, the linguistic response "This is water" will be permitted only when certain, but finitely many, background experiences exist in addition to water-type experiences. In general, background experiences will be experiences of what philosophers often call "standard conditions of observation".

It should not be thought that because objective state-describing responses are allowed by linguistic rules these responses must be true. From the fact that objective state-descriptions are allowed by linguistic rules, given certain families of experiences, it only follows that the objective state-descriptions are justified. It is possible for an assertion, A, to be justified by certain experiences and yet be false. This could happen if certain other experiences occurred which justified some objective statedescription incompatible with A. If the vast majority of our experiences justify a description incompatible with A, then we conclude that A is false, in spite of the fact that A was initially justified by certain of our experiences. On the other hand, if the vast majority of our experiences continue to justify the assertion that A, then we conclude that A is true.

Perhaps an example would be useful here. A rule which would permit the objective state description "This is water" might be something like the

the following: We may say "This is water" if most of the following sets of experiences occur:  $E_1E_2...E_n$ . Thus if a person has most of the experiences  $E_1...E_n$ , that person will be justified in claiming "This is water." Now it may be that each member of the set  $E_1...E_n$  is regularly correlated with other experiences, e.g.,  $E_1$  may be correlated with  $F_1,F_2$ , and  $E_2$  may be correlated with  $G_1,G_2$ , and so on. And it may be that this total set of correlated experiences,  $F_1,F_2,G_1,G_2...M_1,M_2$  is considered relevant to the question whether something is water. So the truth of the statement "This is water" may depend in a given case, upon whether this further set of correlated experiences can be made to occur. If this further set of correlated experiences never occurs it may be appropriate to override the initial judgement "This is water", and to substitute the judgement "This is sulfuric acid". This would not mean, however, that the initial judgement "This is water" was unjustified.

(The example just given is merely speculation on my part. I do not want to commit myself to the particular structure of this example. I merely wish to suggest a possible way in which a justified, rule-guided response to a set of experiences may be overridden by other experiences.)

In the example just described I say "The truth of the statement "This is water" may depend, in a given case, upon whether this further set of correlated experiences can be made to occur." There is, in fact, a difficult and interesting problem about what makes a statement true, as opposed to merely justified. Some claim that to say a particular objective statedescription is true means that the vast majority of our experiences justify, and will continue to justify that state-description, and not its denial. This view is attractive because it supports the view that we in some way construct

our world from our experiences, and there can be no meaningful questions about the objective state of the world which cannot be reduced to questions about possible experiences. On the other hand, it does not seem to do justice to our concept of other people to say that they are constructions of our experiences. I do not wish to take a stand on this controversy here. I bring up the issue to point out that my account of justification can stand whether (a) we think the claim - that assertions which the vast majority of our experiences justify are true - expresses an analysis of our concept of truth, or (b) whether we think this claim expresses some transcendent or metaphysical truth.

Someone might object that my explanation of how some of our experiential beliefs are justified presupposes the existence of an objective world which causes, in part, certain beliefs in us. For example, I claim that we will strongly favor a description which the vast majority of our experiences justify. This seems to imply that usually the vast majority of our experiences will support a particular belief, and that the situation will not often arise where our experiences do not strongly support either a particular assertion or its denial. The question arises whether we are entitled to make this assumption.

To answer this objection: First, I do not think I need to justify assumptions which I make in my explanation of how experiential beliefs are justified. I am offering a theory, and the theory is confirmed to the extent to which it explains what it is intended to explain. Second, within the context of the present discussion it seems odd to call into question the assumption that there is an objective world, and that the vast majority of our experiences will support certain beliefs and not other beliefs. In fact it

seems odd in any theoretical discussion to question the existence of an objective reality. If we do not assume the existence of an objective reality, then what are we theorizing about?

- 1 Karl Popper, The Logic of Scientific Discovery, London, 1968.
- <sup>2</sup> <u>Ibid</u>., p. 103.
- <sup>3</sup> <u>Ibid</u>., p. 106.

### Chapter 6

### Analytic Truth and Necessary Truth

This chapter will deal with the question whether all necessary truths are analytic. One way of attacking this question would be to examine some definition of analytic truth, for example, my own definition, and see whether anything about the definition of analytic truth would enable us to decide the question whether all necessary truths are analytic. I will follow this procedure here.

Recall my definition of analytic truth. According to that definition a sentence is analytic if and only if it is a logical consequence of a correct concept analysis. Now, depending on how we construe the expression 'logical consequence of a correct concept analysis' we may or may not have a ready answer to the question whether all necessary truths are analytic. Most people construe 'logical consequence' in such a way that any conclusion which necessarily follows from certain premises is a logical consequence of those premises. Many logicians also hold that every necessary truth follows from any premises whatsoever, since, if the conclusion of an argument is necessarily true, then it is impossible that the premises should be true and the conclusion false. Given the truth of my claim that 'analytic' applies to any logical consequence of a correct concept analysis, it is likely that these logicians would want to claim that every necessary truth is analytic, since they hold that every necessary truth is a logical consequence of any concept analysis.

Some logicians, however, construe 'logical consequence' more narrowly. And if we do construe 'logical consequence' in a more narrow way the question whether all necessary truths are analytic may be much more difficult to answer. But, in a way, the disagreement over how we should construe the

expression 'logical consequence' may seem idle. After all, what does it matter how we construe this expression as long as we see the consequences of construing the expression each way? The answer to this question is this: I think it is interesting to consider whether there is a more interesting connection between concept analyses and the set of all necessary truths than the mere fact that necessary truths cannot be false at the same time that a concept analysis is true. It would be interesting to see whether we can find a sense for 'logical consequence' in which it would be true to say that a given necessary truth is a logical consequence of one concept analysis, but not another. So let us consider whether we can find such a sense for 'logical consequence'.

Most logicians claim that we can distinguish informal valid consequences from formal valid consequences (a formally valid consequence being one which follows from the premises of a given argument purely in virtue of the form of the argument). These logicians would regard "x is a triangle; therefore, x has three sides" as an informal inference, on the grounds that although this inference is valid it is not valid in virtue of its logical form. There is a problem, however, about how to distinguish those arguments which are formally valid from other arguments. Usually people try to make this distinction by making another distinction, that between logical and non-logical constants. Formally valid inferences are then defined as those in which only logical constants occur essentially. (A term T is said to occur essentially in a valid argument if and only if there is a possible replacement of T by a different term which renders the argument invalid.) Likewise formal logical truths are sometimes defined as true sentences in which only logical constants

occur essentially. (It is commonly recognized that formally valid arguments and formal logical implications are related in the following way: If a conditional is formed which has as its antecedent the conjunction of the premises of a formally valid argument, and as its consequent the conclusion of that argument, then that conditional is a formal logical truth. Such a conditional is called the corresponding conditional of the argument. In general if an argument is valid then its corresponding conditional is necessarily true. In what follows I shall sometimes be discussing the validity of an argument and sometimes the necessity of a corresponding conditional. It should be understood that usually when I discuss validity in this chapter analogous remarks could be made about necessity, and vice-versa.)

Now if we had an adequate definition of 'logical constant' we could make use of the definitions of 'formally valid' and 'formally true' and then we could restrict the class of analytic truths to those which are formally valid consequences of concept analyses. But do we have an adequate definition of 'logical constant'? There is an excellent discussion of this question by Arthur Pap in Ch. 6 of <u>Semantics and Necessary Truth</u>.<sup>1</sup> Pap concludes that there is no important distinction between inferences which are formally valid and those which are informally valid. I will briefly explain how Pap reaches this conclusion.

Pap first rejects attempts to define 'logical constant' by complete enumeration of instances. It is of no theoretical interest to be given a list of logical constants which includes sentential connectives, the existential and universal quantifiers, the identity sign, set membership sign, but which excludes things like 'is larger than', 'is the father of', 'is round', etc., without being told why members of the first list are logical constants while members of the second list are not. It is even likely that people would

disagree about whether certain terms, e.g., the set membership sign, belong in a complete enumeration of logical constants.

Next Pap considers some proposed definitions of 'logical constant' and finds these inadequate. I will discuss only the most plausible definitions. 1. One proposed definition is the following:

"A term is a logical constant if and only if it occurs essentially in some necessary implication." Pap easily produces a counter-example to this proposal, namely: The word 'triangle' occurs essentially in the inference schema "x is a triangle; therefore, x has three sides", although almost no one wants to say that 'triangle' is a logical constant.

2. A modification of the above proposal is also considered by Pap. It is this: "A term is a logical constant if and only if it occurs essentially in every necessary implication in which it occurs".<sup>2</sup> Pap attributes this definition to Reichenbach. He rejects it for the following reason. In the necessary implication  $((p \cdot q) \vee (q \cdot r)) \supset ((q \cdot r) \vee (p \cdot q))$  the occurrence of the conjunctive sign is inessential (the necessity of the implication requires only the commutability of disjunction). But almost everyone, including Reichenbach, admits that the conjunctive sign is a logical constant. So Reichenbach's criterion is too narrow; it excludes something which is a logical constant.

Pap's counter-example to Reichenbach's definition makes use of the fact that the occurrence of any sentential component in a tautology is inessential. Consequently, if a logical constant is a part of such a sentential component its occurrence is inessential. I suggest that we avoid Pap's counter-example by modifying Reichenbach's proposal as follows: R' "L is a logical constant if and only if L occurs essentially in every valid argument

in which it occurs, excluding those arguments where L occurs in a sentence whose occurrence in the valid argument is also inessential." This modification avoids Pap's counter-example. In Pap's counter-example the conjunctive sign '.' occurred in a valid inference, but it also occurred in a sentence whose occurrence in that inference was inessential. Also, it seems clear that the modified criterion, R', would exclude something like 'triangle' from the class of logical constants. For in the following argument the word 'triangle' occurs inessentially within a sentence whose occurrence in the argument is essential:

A triangle exists. Therefore, something exists.

This modification, R', looks promising at first, but there are difficulties with it. For example, consider the following inference:

Some doors are not rotten. Therefore, something is a door. In this valid inference the occurrence of 'not' is inessential. And <u>prima</u> <u>facie</u> the occurrence of 'not' is within a sentence whose occurrence within the argument is essential. So it seems that according to the modified proposal 'not' would not be a logical constant.

Someone might try to meet this difficulty by suggesting that we concern ourselves only with formal representations of valid arguments. For if we symbolize the argument just considered we get:  $(\exists x)(Dx \cdot -Rx)$  Therefore  $(\exists x)(Dx)$ . In this formalization the occurrence of '-Rx', which we may read as "x is not rotten" is inessential. And so, the occurrence of '-' (the negation sign) within this formalized argument is within a sentential component whose occurrence within the valid argument is inessential. Consequently, if we concern ourselves only with formalizations of ordinary language arguments then the argument I have produced would not create a

problem for the modified criterion of logical constants.

However, there is something odd about the suggestion just considered, it is this. We are attempting to find a criterion for saying whether something is a logical constant. Given this goal it is odd to include in the statement of our criterion the requirement that we consider only formalized arguments. For the question arises, what formalizations of ordinary language arguments are we to count as acceptable? Not every formal language is a system of <u>valid</u> inferences, and not every translation of a valid ordinary language argument into a formal language preserves validity. For example, we could translate the inference, "Some doors are not rotten. Therefore, something is a door." into some sentential calculus in the following way: "p, therefore q" (where 'p' and 'q' are sentential constants).

Clearly this formalization is unacceptable for purposes of applying the criterion under consideration. We might avoid this problem by restricting the criterion in question to formalizations which preserve the validity of ordinary language arguments, but this restriction raises two serious problems. They are: (a) all existing formalizations of acceptable logical systems presuppose a prior inventory of logical constants. It would hardly make sense to advance a criterion for logical constants which presupposes a prior inventory of logical constants. (b) If we count "T is a triangle, therefore, T has three sides" as a valid argument in ordinary language, then it is just false that any existing formalized system preserves the validity of <u>all</u> ordinary language arguments. At best, a logistic system such as the <u>predicate</u> calculus preserves the validity of those ordinary language arguments which are, in some sense, valid in virtue of form. Now we might construe the expression

'ordinary language argument which is valid in virtue of form' as referring to those arguments which can successfully be translated into a formal logistic system, but this raises the problem cited in (a).

On the other hand we might try to clarify the notion of an "ordinary language argument which is valid in virtue of form" by compiling a list of ordinary language logical constants. But this alternative is fruitless because it again raises the question "What is a logical constant?"

In view of the problems just cited, I doubt that there is any hope of salvaging the criterion for logical constants which appeals to any <u>precisely</u> defined notion of a formalized valid argument.

Following the next three paragraphs I will offer a <u>vague</u> account of logical constants and formal systems which bypasses these problems. But my <u>vague</u> account of formally valid arguments does not meet the standards of clarification which Reichenbach and Pap have been applying.

Pap discusses another way in which Reichenbach tries to distinguish between logical and non-logical terms, and that is as follows: non-logical terms are denotative terms - terms which denote objects, properties, etc., whereas logical terms do not denote anything and cannot be reduced or defined in terms of denotative terms.

Pap easily produces a counter-example to this definition by pointing out that we could construe the logical constant 'or' as denoting a two-place relational property. And in general we could construe all the sentential connectives as denoting truth-functional relational properties. Thus, the expression 'p or q' could be construed as saying that the propositions "p" and "q" satisfy a particular truth-functional relationship, i.e., the one denoted by 'or'. If it is objected that we could not construe existing truth-

functional connectives as denoting relational properties, because they do not have the appropriate meaning, then we could reply that it would be easy to construct a propositional calculus C, isomorphic to some existing propositional calculus D, such that terms in C which denote relational properties correspond to terms in D which are truth-functional connectives. That would show that we could construct a logistic system in which the logical constants denoted truth-functional relationships. Such logical constants would be denotative terms, and, in fact, if we count the set membership sign as a logical constant, then some logical constants now denote relationships. For the set membership sign denotes the relational property of set membership.

Perhaps even the quantified '(x)' denotes a relation which holds between a formula and all objects in the universe of discourse, namely, the relationship "satisfiable by". It is certainly arguable, at least, that '(x)' does denote this relation - at least as arguable as that words like 'even' and 'as' denote (I assume Reichenbach would want to say that 'even' and 'as' are not logical terms, i.e., they do denote.)

Looking back, we see that all the proposed definitions of 'logical constants' we have considered have serious problems, even when modified to meet initial objections. I think it unlikely that any clear distinction between logical and non-logical constants can be drawn. Perhaps the best we can do is give a vague account of the difference between logical and nonlogical constants. We might say that logical constants are terms which most frequently have essential occurrences in valid inferences. Or perhaps the following suggestion is better. Those expressions are logical constants which may successfully be taken as constants in a very general theory of valid

inference. If by treating a given set of terms as constants we can construct a formal system of inferences, in a fairly economical way, then we may consider that set of terms logical constants. Formally valid inferences may then be defined as those in which only logical constants occur essentially. Or, more directly, we might bypass logical constants and define formally valid inferences as those inferences which could occur in a very general and fairly economical system of valid inference.

If we accept this account of formal inferences, we should note two things, (a) the distinction between formally valid and other inferences is left vague. It is left unclear, for example, whether the inference "A is larger than B. Therefore, B is not larger than A" is formally valid or not. (b) We have defined formally valid inferences as a subset of valid inferences, and we have not attempted to define 'validity' (except in terms of necessity). This is important, because if there is no <u>satisfactory</u> account of formal validity and formal deduction which does not presuppose the concepts of validity and necessity, then my account of analyticity can certainly <u>not</u> be taken as an explanation of the concepts of necessity and validity, since my account of analyticity is given partly in terms of logical consequences, which can only be explained in terms of validity and necessity in general.

Now some philosophers, e.g., Quine, would reject the conclusion that our notions of logical consequence, logical truth, and formal validity can only be explained in terms of validity and necessity in general. In his article "Mr. Strawson on Logical Theory" Quine caustically attacks Strawson for having founded a logical theory on "too soft and friable a keystone" (analyticity and entailment). Quine would object with equal strength if

Strawson had founded his system upon the concepts of necessity and validity. (In fact Strawson was not distinguishing between necessity and analyticity in his book on logical theory). Quine believes that we only understand validity and necessity insofar as we understand formal validity and logical 'Formal validity' and 'logical truth' should be defined, according truth. to Quine, in terms of "statement forms which are logical, in the sense of containing no constants beyond logical vocabulary, and (extensionally) valid, in the sense that all statements exemplifying the form in question are true." $^3$ Quine admits that "logical vocabulary is specified ... only by enumeration", and admits further that this enumeration is apparently arbitrary.<sup>4</sup> So Ouine does not really come to terms with the problem of defining 'logical constant'. Perhaps Quine thinks it unnecessary to define 'logical constant', or perhaps he thinks a sufficient reason for taking certain terms as logical constants is that they have always been recognized as such by logicians. In any case, Quine has no answer to the conclusion I have so far derived, namely, that 'logical constant' can only be defined in terms of the general concepts of validity or necessity.

But let us temporarily suppose, for the sake of argument, that there could be a correct account of logical constants which did not presuppose an understanding of validity or necessity. Then it would be possible to define logical truths in the way Quine wants to do, namely, as those truths which remain true under all reinterpretations of non-logical constants. What I wish to point out, is that any <u>application</u> of this definition of logical truth presupposes prior intuitions about what inferences are valid and what truths are necessary. For the question arises, how could we ever decide, for

example, that all reinterpretations of non-logical constants in the law of non-contradiction, -(p.-p), leave the truth of the law unaffected? Since the number of reinterpretations of 'p' are infinite, we could not test the law under every possible reinterpretation. Furthermore, we could not prove that all reinterpretations of the law do not affect the truth value of the law without presupposing the validity of some other law of logic. We cannot get away from the fact that we cannot establish formal truth of formal validity without presupposing that we know, a priori, that certain truths are necessary and that certain inferences are valid. And it is silly to pretend that we in any way prove the law of non-contradiction is logically true. The way we know that all occurrences of non-logical constants in the law of noncontradiction are inessential is that we can see that the law could not be false, that is, the law is necessarily true. Because of this I think that Quine is wrong when he claims that the concept of logical truth is on a much firmer footing than the concepts of validity and necessity. We can say nothing about what truths are formally true and what inferences are formally valid without appealing in a general way to intuitions of necessity and validity.

I have been arguing that the concepts of validity and necessity are more fundamental than the concepts of formal validity and logical truth. Both Pap and I have serious doubts whether we can even distinguish formal reasoning and formal truth from other kinds of necessary truth (unless my suggestion that we draw the distinction in terms of those inferences which can be systematized is correct). Given this position someone might ask, "What is your account of validity and necessity? It would be pointless to define

necessary truths as the class of all analytic truths since your definition of 'analytic truth' involves the concept of logical consequence, which involves an understanding of validity and necessity. So how do you define validity and necessity?" Answer, I don't. I think validity and necessity are concepts we learn ostensively. Someone says, "If it rains, then the spinach will grow. It rains, so it necessarily follows that the spinach will grow." We get the idea of necessity and validity from this and similiar cases. We can get the idea of necessity from playing games with rules. "You can't do that, it breaks the rules. If you abide by the rules you can only make certain moves." We use the modals 'can' and 'can't' all the time, often to denote absolute possibility and impossibility.

Perhaps we are now in a position to examine the question whether all necessary truths are analytic. I have defined analytic truths as those truths which are logical consequences of a correct concept analysis. We have reason to believe that it is to a large extent arbitrary whether we call a valid inference a formally valid inference. Consequently, there may be no point in distinguishing between formally valid consequences of concept analyses and any other valid consequences of concept analyses, at least for present purposes. When we say that analytic truths are any logical consequences of concept analyses we could mean that analytic truths are any valid consequences of concept analyses. This has the result of making absolutely every necessary truth analytic, provided we accept the principle that an argument is valid if and only if it is impossible that the premises are true while the conclusion is false. Unfortunately for our intuitions, this has the result that the necessary truth which is Goedel's incompleteness theorem is a logical con-

sequence of the concept analysis of "brother".

We might try to avoid this paradoxical result by restricting logical consequences to those consequences of valid inferences which inferences can be readily systematized or included in a general theory of inference. This would rule out rules of inference like, "From any premise you may infer Goedel's theorem". But what about the rule of inference which says that we may infer any necessary truth from any premise whatsoever? That is a valid rule of inference which is of general form and which is included in some systems of modal logic. If we accept that rule of inference, then we are once again stuck with paradoxical results. (For example, within a particular formal system which included this rule of inference we might derive the result that Goedel's theorem is a necessary truth. Then by applying this rule we derive the result that Goedel's theorem is a logical consequence of "2+2=4".) However, I can think of no reason, other than ad hoc reasons, for excluding this as a rule of inference for our purposes. So why not accept the paradoxical result that every necessary truth is a logical consequence of any concept analysis, and accept the corollary that every necessary truth is analytic? If the reader finds this result too counterintuitive to accept, he or she is welcome to construe 'logical consequence' as narrowly as seems appropriate. We could construe 'logical consequence' to mean logical consequence in the predicate calculus, for example. This might have the effect of making many necessary truths, e.g., truths of arithmetic, nonanalytic (since set theory may also be required to derive the truths of arithmetic from a set of definitions). It does not much matter which convention we adopt as long as we realize that we are adopting conventions. If we

construe 'logical consequence' very broadly we will have the Leibnizian result that all necessary truths are analytic and vice-versa. If we construe 'logical consequence' very narrowly we will have the Kantian result that not all necessary truths are analytic. (I do not mean to imply that either Leibniz or Kant arrived at these results by the kind of reasoning I have been presenting.)

## Chapter 6

<sup>1</sup> A. Pap, <u>Semantics and Necessary Truth</u>, New Haven, 1958.

<sup>2</sup> <u>Ibid</u>., p. 136.

<sup>3</sup> Quine, <u>The Ways of Paradox</u>, pp. 138-9.

<sup>4</sup> <u>Ibid</u>., p. 139.
Chapter 7

Convention and Necessary Truth

.

In the preceding chapter I examined the question whether all necessary truths are analytic. The conclusion reached was that all necessary truths are analytic, provided we construe the concept of logical consequence very broadly. If, on the other hand, we construe the concept of logical consequence in a more narrow, and admittedly arbitrary way, then not all necessary truths are analytic. In this chapter I will consider the question whether all necessary truths are the result of linguistic convention. Some philosophers have identified the thesis that all necessary truths are analytic with the conventionalist thesis that all necessary truths are the result of In what follows I hope to show that these two issues are distinct, convention. and that necessary truths are not the result of convention - in any interesting sense. Of course, it is always open to someone to stipulate a use of 'result of convention' which is identical with my use of 'analytic'. In that case my answer to the question whether all necessary truths are the result of convention would be the same as my answer to the question whether all necessary truths are analytic. But, as I will now argue, it would be very misleading to establish such a use for 'result of convention'. (a) Suppose someone claims that a necessary truth (or valid inference) results from our linguistic conventions if and only if that necessary truth (or the corresponding conditional of that valid inference) is a valid consequence of a description of the linguistic rules guiding the use of some word or words in our language. (Wherever possible in the following discussion I will refer to only one of the pair, valid inference/necessary conditional, it being understood that analogous remarks could be made about the remaining member of this pair.) It would be natural to construe this claim as implying that necessity is in some way

created by linguistic convention (otherwise, why say that necessary truths result from linguistic convention?). Now suppose that we could deduce some necessary implication from a description of the rule guiding the use of the English word 'or'. For example, from the rule "the word 'or' may be inserted between any two sentences p and q, provided one of the pair is true" we may validly derive fif p, then p or  $q^2$ , which is the corresponding conditional of  $r_p$ , therefore, p or  $q^2$ . But from the fact that we could perform this deduction we cannot infer that our linguistic rules create logical necessity or logical validity. For, assuming it makes sense to talk of "creating necessity", it is not the linguistic rules alone that create logical necessity (or validity). Rather our linguistic rules together with certain valid rules of inference create logical necessity (validity). But it is absurd to suppose that logical necessity is created in part by logical validity, since a valid inference is just one where the conclusion necessarily follows from the premises.

The point can be more clearly and less metaphorically put as follows: logical truths and logically valid rules of inference are not identical with our linguistic rules. We may stipulate what our linguistic rules are, but we may not stipulate what the consequences of our linguistic rules are. The consequences of correctly following a given set of rules are limited by the rules themselves. In a certain sense it is true that we can stipulate what rule-guided effects will occur by deciding what rules to follow, but once we have decided to follow a given set of rules we can no longer stipulate the consequences of following those rules. In this sense we may not stipulate logical truth or validity, and in this sense it is misleading to say that logical necessity results from linguistic convention.

Michael Dummett makes this argument in a slightly different (and less detailed) form when he discusses what he calls modified conventionalism. Modified conventionalism is the view that although some necessary truths are direct registers of convention, others are "more or less remote <u>consequences</u> of conventions".<sup>1</sup> Dummett's criticism of this view is, "It appears that if we adopt the conventions registered by the axioms, together with those registered by the principles of inference, then we <u>must</u> adhere to the way of talking embodied in the theorem, and <u>this</u> necessity must be one imposed upon us, one that we meet with. It cannot itself express the adoption of a convention."<sup>2</sup>

Against Dummett's argument and my argument one might object as follows: It may be true that certain necessary truths are merely logical consequences of conventions, and not themselves conventions, but these truths can still be <u>explained</u> in terms of conventions. For to explain x in terms of y it surely suffices to show that x can logically be deduced from y. When we say that a necessary truth is the result of convention we mean only that the necessary truth is a logical result of convention, or that it can be logically explained in terms of convention.

Consider another example, suppose the rule guiding our use of 'sister' is that we may apply this word to all and only female siblings. Surely, the citing of this rule would explain why everthing which can correctly be called 'sister' is female, and that explains the truth of "All sisters are female". To be sure this explanation involves the use of logic any explanation does.

So far so good. But what shall we say when we apply these considerations to all logically necessary truths. If every explanation presupposes the use of logic, or the making of valid inferences, then in a sense we can

explain validity and necessity in terms of convention, but if we say this we should also note the oddity of saying that convention is the source of validity and necessity. If there were no validity and necessity there would be no logical results and no explanation. The problem is whether we can explain the existence of something which is a precondition of all explanation. I think we cannot. (Jonathan Bennett discusses this problem briefly in "On Being Forced to A Conclusion".<sup>3</sup> He concludes that <u>either</u> conventions explain the existence of logic, or else the existence of logic cannot be explained at all. I choose the latter alternative.)

(b) I turn now to consider another version of conventionalism -one which Quine discusses in "Truth by Convention".<sup>4</sup> In that article (which contains much that Quine later rejects in "Two Dogmas of Empiricism") Quine gives an argument similar to the one I presented in (a). He points out that the logical consequences of stipulative definitions will be true by convention only if all of logic is true by convention. He then attempts to show how we might construe all of logic as true by convention. The method he describes is commonly known as implicit definition. Briefly, the method is this: "We take one of the many formalizations of the predicate calculus; one in which the logical vocabulary is reduced to a few primitives and in which the basic axioms and rules of inference have been kept to a minimum. We then treat the formal system as an uninterpreted calculus (completely devoid of meaning). Next we stipulate that the basic axioms and rules of inference of the system are to be taken as true. We do not thereby predicate truth of the axioms, that would presuppose that the axioms had meaning. Rather we establish a use for the axioms (and implicitly for their parts) by describing the circumstances

in which they may be counted as true. We count logical axioms as true in any circumstances whatsoever. (More accurately, our formal system contains axiom schemas and inference schemas, and we stipulate that all substitution instances of these schemas are to be counted as true and valid, respectively).

This appears to solve the problem of how logical necessity and validity result from convention directly. Logical axioms and inferences are true and valid because we use them as if they were true and valid, and it is the use of an expression which determines its meaning. Our linguistic behaviour is such that we let the logical particles have any meaning which preserves the validity of our inferences. When we are dealing with a formal system, the stipulations as to which axioms are true and which inferences are valid are made explicit and are verbalized. In ordinary language, however, these stipulations are not made explicit. What justifies us in taking the formal system as a model for what happens in ordinary language is the fact that there is a correspondence between sentences and inferences which are accepted within the formal system and sentences and inferences which we regard as true and valid in ordinary language."

There are numerous problems with the "implicit definition" account of the origin of logical necessity just described. Here are some of them (those due to Quine are so indicated).

1. (Quine) Absolutely any body of doctrine can be rendered true by definition if we follow the method we have considered. For example, we could stipulate that the axioms of physics are true by letting the primitive terms occurring in those axioms take on the required sense. As in the case of logic, the result of formalizing the axioms of physics and stipulating their truth would

preserve our ordinary beliefs about which sentences of physics are true, and which are false. Unfortunately, such a procedure would not preserve our ordinary beliefs about what sentences are necessary and what sentences are contingent. Therefore, those who hold that the necessity of logic and mathematics derives from implicit definition must explain why we conventionally treat the axioms of mathematics and logic as true, but do not so treat the axioms of physics and every other empirical science.

Quine suggests an explanation, namely, that we treat the axioms of logic and mathematics as true by convention, in contrast with physical theories, because of their a priori nature or because they are very deep in our conceptual scheme. Quine's explanation is unsatisfactory, however. The a priori nature of logic and mathematics in ordinary language may explain why we treat certain formulations of these disciplines as conventionally true, but it cannot explain why we treat these disciplines as conventionally true when they are expressed in unformalized ordinary language. We could only claim that the ordinary language expressions of logic and mathematics were a priori if they had some meaning. But on the account we are considering our decision to treat logic and mathematics as true by convention is what endows these disciplines with meaning. Since meaningless theories are not in any conceptual scheme, we can hardly decide to treat a meaningless theory as conventionally true on the grounds that the meaningless theory is a priori. 2. (Quine) When we stipulate that all substitution instances of the basic axiom schema and inference schema of some logical theory are to be counted as true we do so in order to render all logic true by convention. But from

the fact that all substitution instances of the basic axiom schema are true

we can derive the result that a particular substitution instance is true only by inference. This inference is not itself an inference made within the system under consideration, since the system under consideration contains only axiom and inference schema - not particular inferences. Hence any particular substitution instance we can name will not be true purely by convention. It will be the joint result of convention and inference. We might try to avoid this consequence by stipulating that each and every substitution instance of a particular schema is true. But then we are faced with an infinite task.

3. At this point we could repeat Dummett's objection to modified conventionalism, since that is the kind of conventionalism Quine describes. Dummett's objection, remember, is that even granting that all the axioms and rules of inference of a system register conventions to talk a certain way, we still cannot explain why we <u>must</u> accept a certain theorem given the axioms and rules of inference, unless we presuppose the validity of certain logical inferences. Conventions, by themselves, cannot explain their own consequences. We can see that Dummett's objection is related to Quine's objection at (2), insofar as both objections point out that convention, by itself, does not suffice to get us moving, logically speaking. Logic is also required.

4. Quine also makes the point that even the verbal formulation of our adoption of the axioms and rules of inference as conventions presupposes the use of logical vocabulary, i.e., the very idioms which we are purporting to conventionally define. It seems, therefore, that all logic cannot be true by convention if we require logical vocabulary, and consequently logic, to even formulate conventions.

Quine suggests a way of avoiding this objection. He points out that in ordinary language the convention to adopt certain ways of speaking need never be verbalized. Indeed, if all linguistic conventions had to be introduced by explicit verbal agreement, language could never get off the ground, since we would need a language to start a language. Consequently, there is no need to suppose that the conventions to adopt certain axioms and rules of inference in ordinary language ever were, or ever need be verbalized. Thus the problem is avoided. That our linguistic behaviour is guided by convention, someone could maintain, is shown by the regularity our linguistic behaviour exhibits, quite apart from explicit verbal conventions.

Quine does not reject this solution (not in "Truth by Convention" at least). He does point out, however, that the idea that there might be unverbalized, unexplicit conventions which are manifested by behavioural regularity is in need of clarification. I have argued (in an earlier chapter) that such a clarification has been given by David Lewis. I am referring to the criteria for rule-guidedness which I described in chapter 4.

(c) The next version of conventionalism I wish to consider is as follows. "Logical necessity and validity result from linguistic conventions in the sense that if any person denies a necessary truth, S, we are justified in concluding that the person does not know the meaning of some word or words in S, i.e., does not know all the linguistic conventions guiding the use of all or part of S." For example, if someone asserts that "Grass is green and the sky is blue" is true, but denies that "Grass is green" is true, we would have to conclude that the person does not know how to use the word 'and', i.e., the person's use of 'and' is not guided by the proper rules. If the person

did not believe that "Grass is green" is true, then he or she is not permitted by the rules guiding the use of 'and' to conjoin "Grass is green" with any other sentence. Roughly, the rule guiding our use of 'and' is that we may insert 'and' only between sentences that we believe are true.

Although the example we have just considered is a convincing one, a lot more is required to demonstrate that whenever a person denies a necessary truth or valid inference that person displays a misunderstanding of some piece of language. For the example we have just considered is much too special a case. There seems to be no room for mistake, other than a linguistic one, p and q, one asserts p and q and denies p. In this sense when, for some we can explain the validity of the inference from r p and q r lo p by describing the rules guiding our use of the word 'and'. But other valid inferences are more complex and cannot be so closely linked with linguistic rules governing the use of logical particles. Because one can know all the linguistic rules guiding our use of logical vocabulary without seeing all the consequences of these rules, one can fail to see that validity of a complex inference without thereby displaying a misunderstanding of any logical vocabulary. Likewise, one can know all the rules of chess without seeing all the consequences of these rules in a particular chess board situation. This is not a new observation, or a very controversial observation, but I think it would be interesting to see why it is true. To this end I will try to explain in more detail how, for example, it is possible for someone to deny even so simple a necessary truth as that 7+5=12 without thereby displaying a misunderstanding of any arithmetical symbol or ignorance of any linguistic conventions.

My explanation is as follows: Suppose we take Frege's account of

numbers to be correct. On this supposition numbers are sets. The number zero, for example, can be defined as the set of all objects which are not self-identical. The number one can be defined as the set whose only member is the number zero. The number two can be defined as the set whose only members are the numbers one and zero. And so on. Each number after zero can be construed as the set S whose only members consist of all the numbers which occur earlier in the series of sets than S. On this account, and indeed on every other theory of numbers which is adequate, the integers constitute an ordered, linear series of objects. This means that we can safely view the whole numbers as a series of points on a line, some numbers appearing later in the series than others. Now it is a convention that we use '12' to denote a number which occurs later in the number series than the number we denote by '5'. But the fact that the number so denoted occurs later in the series is no convention; it is something given to us. Thus even the simple arithmetic truth that 12 is greater than 5 is neither a convention, nor purely the result of convention. Rather it is the joint result of convention and fact. Consequently, the truth that 12=5+c, where 'c' denotes some positive number, is not purely the result of convention. We all know that c turns out to be the number 7, but that c=7 once had to be discovered; it is not a convention. One way to discover that c=7 is to count the integers in the closed interval from 6 to 12. Now it is possible to make a mistake in counting without thereby demonstrating that one misunderstands any numerals or how to count. As a consequence it is possible to think that 7+5=13, and to deny that 7+5=12, without thereby displaying that in general one does not know how to use numerals or to count.

I have portrayed our discovery that 7+5=12 as a discovery of fact. The question arises, how do I reconcile this view with the common belief that "7+5=12" is an a priori necessary truth? I will try to answer this question.

We can formulate a general rule for adding numbers. To add a number x to a number y start with the immediate successor to x in the number series and count, in order, the integers following x until you arrive at the yth such integer. The yth integer will be the sum of x and y. Now, in a sense, when we sum two numbers according to the rule just given we make an empirical discovery; we are counting some objects. But the objects we are concerned with, i.e., the integers, can be generated according to an exact rule (e.g., construct each number by adding one to its antecedent. The first number is zero.) Integers are also named according to an exact rule, and the summation procedure is purely a rule-guided process. So, in a sense, the result of a particular summation is determined by the rule for generating integers, the naming rule, and by the rule guiding the process of summation. Of course, in a particular case it is always possible to miscount by accidentally omitting an integer or by counting twice. In that sense the result of an individual count is contingent. But, provided we follow the counting rules correctly, the result of an individual count of the integers between 5 and 12 is not contingent.

Analogous remarks can be made concerning logical proofs. Suppose we blindly apply some rules of inference to a set of premises and generate an unexpected conclusion. Each step in our proof is the result of applying a rule (which may implicitly embody a convention) but the conclusion we generate

comes as a surprise to us. And when we discover that our conclusion can be deduced from our original set of premises we discover a fact, pure and simple. It is, of course, contingent that we generated the conclusion we did, but it is not contingent whether the conclusion can be generated by <u>correct</u> application of the rules of inference. When we discover, in a given proof, that a certain conclusion follows logically from particular premises we discover a fact, namely,that the parts of the conclusion-sentence stand in a particular geometrical relation R to the parts of the premise-sentences. But it is not a contingent question, given the valid rules of inference, whether objects which stand in the relation R are related as logical consequence to premise.

There is more to be said. We see that the fact that summation and proof are rule-guided processes does not suffice to explain the source of necessary truth, though it might explain why the results of a particular investigation are determined, given that the investigation is carried out correctly, and given the facts being investigated.

For example, we can explain why someone who counts the integers in the closed interval from 6 to 12 will always count 7 integers, provided the count is correctly made, and given that there are 7 integers in that interval. But how do we explain the fact that there are, and must be, seven integers in that interval? Well, we could point out that unless there were 7 integers in that interval we could not be talking about <u>that interval</u>. Since numbers do not change through time the relations between numbers do not change through time. Consequently, there will always be 7 integers in <u>that</u> interval. Also the criteria for identity of numbers is exact. I think this results from the fact that the rules governing our use of numerals are perfectly exact.

This exactness is possible because numbers do not change through time, and because (as we have noted) the relations between numbers do not change through time.

I think the case is similiar for logical terms. The rules guiding their use are very exact, because two sentences either stand in a particular truth-functional relation or they do not. Consequently, the criteria for identity of logical truths is very exact. Perhaps all necessary truths involve this kind of exactness. This is not very clear, but I think it is suggestive.

(d) I turn now to discuss Dummett's view of Wittgenstein's brand of conventionalism. According to Dummett, Wittgenstein once believed that any truth which we accept as necessary expresses a convention to talk a certain way. For example, the mathematical truth that 7+5=12 is not a consequence of convention, but is itself a convention. 7+5=12 because our criterion for saying that someone has added 7 and 5 correctly is that the result be 12.

At first glance this kind of conventionalism seems very implausible. One is tempted to say that it is not open to us to stipulate the truth of 7+5=12, because the truth or falsity of that statement is already determined by the conventions we have laid down governing addition in general. For example, we might construe the convention governing addition as follows: to obtain the sum a+b simply count b digits past the number a. The bth digit past the number a is the sum of a and b. Thus, to add 7 to 5 we simply count 7 digits past the number 5. The 7th digit we count will be the sum. Now it certainly seems that the sum of 7+5 is determined by this counting procedure, and any convention we lay down dictating what this sum should be runs the

risk of conflicting with the conventions we have already laid down. We could not, for example, stipulate that the sum of 7+5 is 13, because this would conflict with other conventions.

Wittgenstein might reply as follows: In what sense is the sum of 7+5 determined by existing conventions of counting rules? The sum is not determined in the sense that whoever counts 7+5 objects must arrive at the number 12, for it is always possible to miscount. Likewise, the sum is not determined by any proof, since it is always possible that any proof we construct contains some mistake. Both the counting method and other methods of proof are fallible. Consequently, neither method <u>forces</u> any conclusion on us. Since both methods are fallible processes we need some criterion for deciding whether we have counted correctly or carried out the proof correctly. And since, most people do arrive at a count of 12 when counting a collection of 7 objects and 5 objects it is convenient to make the sum, 12, the criterion for a correct summation of 7 and 5. But this is purely a practical matter.

So far we have stated an objection to Wittgenstein's theory and suggested a possible answer on Wittgenstein's behalf. But this answer is inadequate. This is apparent once we consider that we do not need to take the arrival at an orthodox result as the criterion for saying whether the result was correctly deduced. We have independent ways of checking, for example, whether a count was correctly made, or whether a theorem was correctly derived. We merely check to see whether each move was made according to the rules. We can check someone's counting of a set of objects by having the person number the objects as he counts. We then check to make sure that each object was numbered by a different number and in the correct sequence. Of course it is

possible that we all make the same mistake in our checking, but the chances that we would all arrive at the same result by accident are very small. Consequently, if a large number of people check a person's count (proof) and arrive at the same result, we have very good evidence that the count (proof) is correct or incorrect, as the case may be. This, then, constitutes one important objection to the proposed answer.

There are other problems as well. For example, Wittgenstein has no way of explaining why, when most people count 7+5 and arrive at 12, we can find no mistake in their counting, i.e., no place at which they violated a counting rule. And he cannot explain why we can always discover a mistake when a person counts 7+5 and does not get 12. On my account it is easy to explain these matters. Miscounting consists in violating a counting rule; not in arriving at a certain wrong number. The reason people usually get 12 when they count 7 objects and 5 objects is that they usually follow the counting rules correctly (usually we can discover no mistake), and it is an absolute necessity that if one counts 7 objects and 5 objects correctly, then one will have counted 12 things. This necessity cannot itself be explained in terms of other conventions, however, since it will always be true of any conventions we cite that when they are followed certain results must occur.

Wittgenstein would be right to point out that it is always possible to miscount. And there is no verification of the fact that 7+5=12 which does not rest upon some fallible procedure such as counting or proof. In order to reconcile this fact with the fact that we all regard 7+5=12 as an unassailable truth, Wittgenstein might suppose that we in fact treat 7+5=12 as a convention, that is, we will not allow 7+5=12 to be falsified. I think

Wittgenstein's reasoning may take the following form. "The conclusion of any proof is fallible, since any proof may contain a mistake. "7+5=12" is an a priori necessary truth. Conventions are not fallible. Therefore, "7+5=12" expresses a convention." The fallacy in this reasoning, I suggest, lies in equating a priori necessary truths with infallible truths. But, as I have argued in an earlier chapter, there is no absolute infallibility. Any judgement can be mistaken - including the judgement that we are adopting a convention. Wittgenstein would be right to note that if the verification of "7+5=12" is made to rest upon counting, or upon proof of any kind, then that verification is fallible, and the result of that verification comes as a discovery. But I have been arguing that the fact that "7+5=12" must be discovered is compatible with that truth's being a priori, necessary, and nonempirical. What ensures the non-empirical character of this arithmetical proposition is that, although its proof must be discovered or invented by us, that proof must be generated according to rules, if it is to be correct. And if a proof is generated according to the proper rules, then the conclusion will be necessarily true, provided the premises are necessarily true.

In the preceding section I have tried to dispose of one possible argument Wittgenstein might use to support his brand of conventionalism. There is, however, another line of argument which is discussed by Jonathan Bennett in "On Being Forced to a Conclusion"<sup>5</sup>. It relates to the argument just considered in that it also focusses on a problem about proof. What follows is a simplified account of Bennett's reconstruction of Wittgenstein's argument.

As Bennett sees it Wittgenstein was led to reject modified conventionalism in favor of a bolder form of conventionalism, because

Wittgenstein believed that his behavioural theory of meaning (to be explained in a minute) is incompatible with the traditional picture of logical committal and because modified conventionalism requires the truth of the traditional picture of logical committal. Bennett describes the traditional picture of logical committal as that according to which we are absolutely committed to certain conclusions by the adoption of certain premises. (There is no room for choice on the traditional picture of logical committal.) Now it is clear that modified conventionalism does indeed require the truth of the traditional theory of logical committal, since according to modified conventionalism we are absolutely committed to the truth of certain sentences by the conventions of language we adopt. So if there really is an incompatibility between Wittgenstein's behavioural theory of meaning and the traditional theory of logical committal, Wittgenstein was right to reject modified conventionalism. Let us consider whether such an incompatibility does exist.

Bennett describes the behavioural theory of meaning as the view that the only evidence for what a piece of language means is how we use it. Furthermore, "to mean such and such by a noise <u>is</u> just to be disposed to use it in certain ways."<sup>6</sup> Our use of language is not guided by meanings or rules, as if meanings and rules were over and above behaviour. Rather, our linguistic behaviour defines the meanings of words and determines linguistic rules. Linguistic rules do not <u>prescribe</u> linguistic behaviour, rather they <u>describe</u> it. Now the <u>prima facie</u> incompatibility between the behavioural theory of meaning and the traditional picture of logical committal is this. The uttering of certain sounds or the writing of certain signs at one time cannot commit one to uttering or writing any specific thing at another time. There can be

nothing incorrect about a community's using words according to one pattern before time T and according to a different pattern after T. For the two patterns taken together create a pattern which give their words a unitary meaning. Thus there seems to be nothing to prevent a community of people from assenting to a certain set of written premises and denying a conclusion we normally think of as being entailed by those premises. Such linguistic behaviour may be deviant with respect to our linguistic behaviour, but we cannot say the behaviour is wrong. We can conclude that such a deviating community must mean something different by the premises and conclusion than we do, but on the behavioural theory of meaning this is just another way of saying that the deviators' linguistic behaviour is different from ours. We cannot say to people in the deviating community "You are committed to accepting a certain conclusion by the meaning you attach to these premises," because what meaning they attach to the premises is determined, in part, by what conclusion they are willing to draw. We can form inductive hypotheses about other people's future linguistic behaviour, and hence about what they mean by certain words, but there is no necessity about what people's future linguistic behaviour will be like. The possibility that some sub-community of our linguistic community will deviate linguistically from our community cannot be ruled out a priori. This is the source of the problem about saying that other people are logically committed to certain conclusions. Whether the problem can be solved we shall now consider.

In his article "On Being Forced to a Conclusion" Bennett reconciles the behavioural theory of meaning with the traditional picture of logical committal in much the same way that many philosophers of science reconcile everyday scientific practice with the problem of induction. Scientists do not make room within scientific reasoning for the possibility that nature should cease to be lawlike - primarily because science would not be possible in that case. Likewise we do not make room within logic itself for the possibility that some linguistic sub-community will deviate from us significantly in their linguistic behaviour, for if that possibility should be realized in a generalized way communication would breakdown, at least in a limited area. Bennett put the matter thus: "We make no room within the communication-game for the possibility that the game will become unplayable, just as we make no room within science for the possibility that science will cease to be a possible kind of activity."<sup>7</sup>

The question arises, what do we mean by saying "we make no room within logic for the possibility of linguistic deviation"? I take this expression to mean that all proof takes place relative to the assumption that we are not dealing with linguistic deviators. On the behavioural theory of meaning this assumption is equivalent to the assumption that we are dealing with people who mean by words approximately what we mean by them. This agrees with what most philosophers, including those who posit the existence of propositions, have traditionally said. Most philosophers would say that assenting to certain sentences commits one to certain conclusions only on the assumption that one means by the premise-sentences certain things and not other things. The behavioural theory of meaning supports this traditional view, but unlike traditional theories of meaning the behavioural theory of meaning does not in any way <u>explain</u> why people usually draw the same conclusion from the same premises. On the behavioural theory nothing is explained by

saying that people who mean the same things by premises P will usually draw the same conclusion, and nothing is explained by saying that one can avoid being committed to certain conclusions by assigning a different meaning to the premises. This is because, on the behavioural theory, what one means by premises P is determined, in part, by what conclusions one is willing to draw, not vice versa.

So far I have described, in a simplified way, how Bennett reconciles the behavioural theory of meaning with the traditional picture of logical proof. As far as the reconciliation goes, I think it is the best reconciliation that could be given, but I am doubtful whether it is entirely successful. My reasons are as follows.

Bennett's theory does not explain how individuals can know about themselves that they are committed to a certain conclusion. For example, if I accept both "It is raining" (P) and "If it is raining, then the pollution will clear" (Q), I can see that I am committed to the conclusion "The pollution will clear" (R). And, in general, I often know what conclusions I am committed to, and I feel that I must accept these conclusions. Now the behavioural theory of meaning does not explain why we often feel that we <u>must</u> accept a certain conclusion if we are to be fully rational, though it <u>might</u> explain how we could know that we <u>are going</u> to accept a given conclusion. On Bennett's theory we could correctly say that someone <u>must</u> accept conclusion R if he or she accepts premises P and Q <u>and if</u> that person is not a linguistic deviator. But this is not relevant to our present case, since, when a person sees that he or she is committed to R on a certain understanding of P and Q, it is irrelevant to that person whether he or she is a linguistic deviator.

Even linguistic deviators can be committed to certain conclusions on their understanding of the premises. A person can feel and be committed to a conclusion without even knowing whether he or she is a linguistic deviator, or is disposed to be a linguistic deviator. I think we must conclude that it is one's understanding of the premises which determines whether one is committed to a certain conclusion - not one's dispositions to deviate linguistically. (Of course, someone holding the behavioural theory of meaning may claim that to <u>understand</u> a set of premises in a certain way is just to have certain linguistic and other dispositions. But we could hardly explain the fact that we <u>feel compelled</u> to accept certain premises by appealing to our present linguistic dispositions, for a large part of the evidence that we have certain linguistic dispositions is that we feel compelled to accept certain conclusions.)

Now this difficulty could be avoided by a particular modification of Bennett's position. The modification is this. Instead of trying to explain the origin of the feeling of logical committal by talking about linguistic deviators, we recognize that linguistic deviation is irrelevant to whether an individual <u>is</u> or <u>feels</u> committed to a certain conclusion. What is relevant, however, is one's own internal neural state. Presumably, when one learns to speak a language some change occurs in one's brain. (Such a neural change, or something like it, must exist if we are ever to explain, physiologically, how someone passes from a pre-linguistic state to a state of linguistic competence.) Also, we may assume, people who speak the same language have certain neural or brain structures in common. Now, given that an English speaker, say, has certain brain features unique to English speakers, these brain featrues may occasionally cause an English speaker to accept a certain

conclusion once he or she has accepted certain premises. This could account for those cases where people feel they must accept a particular conclusion. People feel they must accept the conclusion because they are physiologically compelled by their neural structure to accept it.

This modification of Bennett's position does explain why people often feel they must accept a certain conclusion. It also can be used to explain a related problem, namely, why people who accept the same premises generally can be made to accept the same conclusion. But this modification does not adequately explain all aspects of the traditional picture of logical committal. For example, suppose a group of English speakers all accepted a set of premises, and suppose also that one day the old laws of brain/neural physiology ceased to hold. It might then be true that the English speakers would no longer feel committed to accept any conclusions, since they would no longer by physiologically compelled to accept any conclusion. But according to the traditional picture of logical committal the group would still be committed to certain conclusions. According to the traditional picture people are committed to conclusions whether or not anyone feels committed. It is difficult to see, therefore, how Bennett could explain the existence of logical committal in such a circumstance (given the modification we are considering). He might deny its existence, but then he would have failed to reconcile the behavioural theory of meaning with the traditional picture of logical committal.

This completes my list of doubts about the success of Benmett's reconciliation of the behavioural theory of meaning with the traditional view of logical committal. I do not claim that the problems I have uncovered dis-

prove Bennett's view, but they are problems which can be avoided altogether if one rejects the behavioural theory of meaning. Very little reflection will show that I have, so far in this thesis, adopted a position inconsistent with the behavioural theory of meaning. For it has been part of my thesis that the regularity which language exhibits is explained by the fact that our use of language is rule-guided - rule-guided in the sense that there are rules which underlie our use of language. In my sense of rule-guided, to say that language is rule-guided is not just to say that language use falls into regular patterns. I believe that the rules which guide our use of language are inferred entities which explain the regularity of language in something like the way in which a computer program explains the behaviour of a computer. Thus I reject the behavioural theory of meaning in favor of the theory which says that language use is rule-guided.

In "On Being Forced to a Conclusion" Bennett says that "The stress of rules, which is legitimate in itself, could mislead us into denying that logical committal is reducible to relations amongst complex sets of noises. Wittgenstein saw this danger, and insisted at length that problems about meanings cannot be settled just by an appeal to rules, because there will always remain the problem of the meanings of the rules."<sup>8</sup> In adopting the theory of rules I have adopted I have implicitly rejected the simple-minded theory of rules which Wittgenstein appears to have in mind. People do not learn the rules of language by being told what the rules are, rather they learn the rules by observing people's behaviour. And the evidence for what linguistic rules a person is following is the person's linguistic behaviour not what the person says about rules. Consequently, if a linguistic subcommunity of our linguistic community starts deviating linguistically with respect to what proofs they accept as logically valid, and if its members cannot be made to admit a mistake, then we have extremely good evidence for saying that their use of some portion of our language is guided by different rules than our use. This remains true regardless of what rules they say they are following, since, although they may say they are following the same verbal rule as we are, they may understand that verbal rule differently from us. (This view seems a natural one - Bennett even seems to adopt this view in his article when he says, "We could show him (a linguistic deviator) that his use of the word <u>required the learning</u> of rules for our use of it, plus, as a sheer addition, the learning of rules for his use of it."<sup>9</sup> The idea that the use of language requires the learning of rules is closely connected with the view that language use is guided and explained by rules, and does not sit happily with the behavioural theory of meaning which Bennett endorses elsewhere in his article.

It has always been a part of the traditional picture of logical committal that premises commit one to a certain conclusion only on a certain interpretation of those premises. On my theory, to interpret premises in a particular way is to allow one's use of those premises to be guided by <u>certain</u> linguistic rules. Also, on my theory, it is clear that we have very good evidence for saying that linguistic deviators are following different linguistic rules from our linguistic group. Consequently, if one adopts the theory of rules I have been defending, it is easy to explain why linguistic deviators do not present a problem for the traditional picture of proof. Linguistic deviators who reject traditional proofs are following deviant linguistic rules, and thus are not committed to conclusions we standardly accept (given

standard linguistic rules). This is one explanatory power of the ruleguidedness theory. There are others. For example, on this theory, the reason why people often feel that they must accept a certain conclusion, is that their linguistic behaviour is guided by certain rules. If they were to deny the conclusion which they feel they must accept, then they would, at some point, be violating a linguistic rule they have learned (without knowing it). This point connects with a problem mentioned earlier, namely, why most members of a linguistic community who accept the same premises can be made to accept the same conclusion. The solution to this problem, I think, is that those people are guided by the same linguistic rules (in the relevant areas). Consequently, those people can be presented with an argument whose individual steps are small enough to enable those people to feel that they must accept the conclusion of each step. And if we can get a person to accept the immediate conclusion of each step of an argument we can get that person to accept the conclusion of the last step of the argument, that is, the conclusion of the entire proof.

A question which may now occur to the reader is, "In what way do rules guide our linguistic behaviour?" or "By what mechanism do rules guide our linguistic behaviour?" This question was answered when I first introduced the rule-guidedness theory, but because the question is important and relevant to the present discussion I repeat my answer here.

I said earlier that linguistic rules guide our linguistic behaviour in something like the way in which a computer program guides a computer. In fact I think linguistic rules constitute a partial program of our brains. This programming could occur if each brain were innately programmed to program

itself further, according to the particular linguistic environment it was placed in, and there is no reason why humans could not do the same thing, at a neural level. It seems plausible that children do something like this when they learn the generative grammar of the language they learn to speak.

Certain neural changes may cause us to use the word as if we were consciously following a certain rule. In such a case it may be appropriate to say that the rule is guiding our use of the word. For example, if (a) a creature's neural structure changes as a result of the creature's having observed some linguistic regularity which holds by convention (in Lewis's sense), and if (b) the linguistic behaviour produced by that neural change is such as would be produced by consciously following a particular rule, then we may say that the linguistic behaviour is guided by the linguistic rule. We may say this because (1) the relevant linguistic behaviour is being guided by a neural change which occurred in order to enable the creature to imitate a linguistic regularity, and (2) the linguistic behaviour, and (3) the linguistic regularity being imitated is a regularity which exists by convention (in Lewis's sense). This, I suggest, is the mechanism by which conventions guide linguistic behaviour.

Now the question arises, what is the advantage to saying that people's linguistic behaviour is guided by rules rather than by neural impulses of a certain kind. The advantages are several. For one thing, (as was pointed out in an earlier chapter) if we suppose linguistic behaviour is rule-guided we can explain why it is appropriate to criticize deviant linguistic behaviour as incorrect. For another thing, we can explain the obvious element of convention involved in using the word 'rain' to denote rain, rather than the word 'snain'.

A third advantage is that we can explain (where Bennett's theory failed to explain) why people would still be committed to a certain conclusion by the adoption of certain premises even if the laws of physiology should cease to hold. To see this consider the following. According to my theory, if the laws of physiology should break down, then our linguistic programming would fail to be carried out, but it would still be true that we were programmed to follow certain rules. Likewise, if a computer breaks down its program may fail to be carried out, but it will remain true that the computer was programmed. Now given the linguistic rules that program our linguistic behaviour, and given that we are programmed to follow these rules (by our linguistic training), and given that we assent to certain sentences, we are committed to certain conclusions, even if the laws of physiology cease to hold. This is because, what conclusions we are committed to are determined by what linguistic rules we have been programmed to follow. The program need not actually be carried out for this committment to exist, because we are committed to those conclusions which we would accept if we did follow the linguistic rules of our mental program.

This ends my discussion of how Bennett attempts to render the behavioural theory of meaning compatible with the traditional picture of logical committment. I conclude that the two theories cannot be made compatible and that Wittgenstein was right to reject the traditional picture of logical committment <u>given his acceptance</u> of the behavioural theory of meaning. However, I think I have given good reasons for rejecting the behavioural theory of meaning in favor of the rule-guidedness theory of

meaning. So I think Wittgenstein was wrong to hold the behavioural theory of meaning, and consequently his reasons for rejecting modified conventionalism were also wrong (Assuming Bennett is correct in thinking that Wittgenstein's behavioural theory of meaning led to his rejection of modified conventionalism.)

## Chapter 7

<sup>1</sup> Michael Dummett, 'Wittgenstein's Philosophy of Mathematics", reprinted in <u>Wittgenstein</u>, ed. by Pitcher, Garden City, N.Y., 1969, p. 424.

<sup>2</sup> <u>Ibid</u>., p. 425.

- <sup>3</sup> Bennett, "On Being Forced to a Conclusion", <u>Proceedings of the Aristotelian</u> Society, Supplementary Vol. 35, 1961.
- <sup>4</sup> Quine, The Ways of Paradox.
- <sup>5</sup> Bennett, Op. Cit.
- <sup>6</sup> Ibid., p. 16.
- <sup>7</sup> Ibid., p. 32.
- <sup>8</sup> Ibid., p. 17.
- <sup>9</sup> Ibid., p. 32.

## SELECTED BIBLIOGRAPHY

- Aune, B., "Is There an Analytic A Priori?", <u>Journal of Philosophy</u>, Volume 60, 1963.
- Ayer, A.J., Language, Truth and Logic, London, 1946.
- Bennett, J.F., (1) "Analytic-Synthetic", reprinted in <u>Necessity</u>, ed. by Sumner and Woods, New York, 1969.
  - (2) "A Myth About Logical Necessity", <u>Analysis</u>, Volume 22, 1961.
  - (3) "On Being Forced to a Conclusion", <u>Proceedings of the</u> <u>Aristotelian Society</u>, Supplementary Volume 35, 1961.
- Bradley, R.D., "Geometry and Necessary Truth", Philosophical Review, Volume 73, 1964.
- Carnap, R., "Meaning and Synonomy in Natural Languages", <u>Philosophical Studies</u>, Volume 6, 1955.
- Casteneda, H-N., "'7+5=12' as a Synthetic Proposition", Philosophy and Phenomenological Research, Volume 21, 1960-61.
- Church, A., "Intensional Isomorphism and Identity of Belief", <u>Philosophical</u> <u>Studies</u>, Volume 5, 1954.
- Dummett, M., 'Wittgenstein's Philosophy of Mathematics", reprinted in Wittgenstein, ed. by Pitcher, Garden City, New York, 1966.
- Gasking, D.A.T., "Mathematics and the World", reprinted in Logic and Language, Second Series, ed. by A. Flew, New York, 1965.
- Grice, H.P., and Strawson, P.F., "In Defense of a Dogma", reprinted in Necessity, ed. by Summer and Woods, New York, 1969.
- Harman, G., "Quine on Meaning and Existence", <u>Review of Metaphysics</u>, Volume XXI, 1967.
- Hintikka, J., "Are Logical Truths Analytic?", reprinted in <u>Necessity</u>, <u>Op. Cit</u>.
- Kant, I., Critique of Pure Reason, translated by N.K. Smith, London, 1933.
- Katz, J.J., (1) "Analyticity and Contradiction in Natural Language", <u>The Structure of Language</u>, ed. by Fodor and Katz, Englewood Cliffs, N.J., 1964.
  - (2) "Some Remarks on Quine on Analyticity", reprinted in <u>Necessity</u>, <u>Op. Cit.</u>

Kneale, W.C., and Kneale, M., The Development of Logic, Oxford, 1962.

- Lewis, D., Convention, Cambridge, 1969.
- Mates, B., "Synonymity", University of California Publications in Philosophy, Volume 25, 1950.
- Pap, A., Semantics and Necessary Truth, New Haven, 1958.

Popper, K.R., The Logic of Scientific Discovery, London, 1968.

- Putnam, H., (1) "The Analytic and the Synthetic", <u>Minnesota Studies in the</u> <u>Philosophy of Science</u>, III, ed. by H. Feigl and G. Maxwell, <u>Minneapolis</u>, 1962.
  - (2) "Some Issues in the Theory of Grammar", <u>Proceedings of the</u> <u>Twe1fth Symposium in Applied Mathematics</u>, American Mathematical Society, Providence, 1961.

Quine, W.V.,	(1) "Truth by Convention", The Ways of Paradox, New York, 1966.
· ·	(2) "Mr. Strawson on Logical Theory", Ibid.
	(3) "Carnap and Logical Truth", Ibid.
	(4) "Two Dogmas of Empiricism", From a Logical Point of View,
	Cambridge, 1953.
	(5) "The Problem of Meaning in Linguistics", Ibid.
	(6) "On a Suggestion of Katz", reprinted in <u>Necessity</u> , <u>Op. Cit</u> .
<b>a a</b>	

Sommers, F., "Meaning Relations and the Analytic", <u>Journal of Philosophy</u>, Volume 60, 1963.

Ziff, P., Semantic Analysis, Ithica, 1960.

## FOOTNOTE INDEX

Footnotes for each Chapter appear on the following pages:

Chapter	1	•	•	•	•	•	•	•	•	•	•	•	•	•	6
Chapter	2	•	•	•	•	•	•	•	•	•	•	•	•	•	35
Chapter	3	•	•	•	•	•		•	•	•	•	•	•	•	45
Chapter	4	•	•	•	•	•	•	•	•	•	•	•	•	•	6 <b>9</b>
Chapter	5		•	•	•	•	•	•	•	•	•	•	•	•	86
Chapter	6	•	•	•	•	•	•	•	•	•	•	•	•	•	102
Chapter	7														132