A DEFENCE OF IMMATERIALISM

A DEFENCE OF IMMATERIALISM

By

HOWARD JAMES SIMMONS, B.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts

McMaster University

January 1983

MASTER OF ARTS (1983) McMASTER UNIVERSITY
(Philosophy) Hamilton, Ontario

TITLE: A Defence of Immaterialism

AUTHOR: Howard James Simmons, B.A. (Oxford University)

SUPERVISOR: Professor N.J. Griffin

NUMBER OF PAGES: vi, 112

ABSTRACT

This thesis is an attempt to defend the somewhat outrageous view that the only things we need consider to be real are experiences. A serious objection to this view is that we seem to talk of things other than experiences, and that such talk is apparently unavoidable. But in Chapter One it is argued that if non-experiential talk can be interpreted in what is called a 'non-representational' way, then the apparent reference to non-experiential entities can be considered illusory. In Chapter Two a method of doing this is outlined. The key concept here is that of the 'acceptability of a sentence relative to a set of experiences.' Roughly, it is proposed that a semantic theory be a system of rules that determines this property for any assertoric sentence and any set of experiences. In Chapter Three the concept of truth is discussed in terms of the ideas already introduced. It is argued that we should recognize two kinds of truth - what are called 'redundancy truth' and 'empirical truth'. Empirical truth is more important for our purposes, but we have to recognize that there may be some sentences that are neither empirically true nor empirically false. Although this might be thought to necessitate a revision of classical logic, it is argued that this is in fact not so.

The last chapter contains an evaluation of the traditional empiricist conception of experience and attempts to defend the following: the epistemological priority of experiences; the withholding of ontological status from physical things; and the rejection of solipsism. There is also a comparison (though only a limited one) between the present theory and the verificationism of M.A.E. Dummett.

ACKNOWLEDGEMENTS

I would like to thank Professor Nicholas Griffin and Professor Neil Wilson (especially the former) for many helpful criticisms and suggestions and also Mrs. Barbara Cameron for doing most of the typing, coping so well with its many horrors.

TABLE OF CONTENTS

CHAPTER ONE: A CLARIFICATION OF THE PROBLEM

Let us begin with an attempted clarification of the problem of realism. Firstly it is obvious that to talk simply of 'realism' is merely shorthand. There are different forms of realism each asserting the existence of a distinct category of objects. There is realism with respect to moral properties, beauty, universals, physical things and so on. In saying that I am defending 'immaterialism', I mean that I am opposing realism with respect to the category of physical things. To ask about the correctness of realism with respect to some category of objects is arguably to ask a verbal variant of the question of whether we should be 'ontologically committed' to that category of objects. But before we try to answer this question for any category of objects at all, it is clear that we must address ourselves to a more fundamental one viz. what <u>is</u> it for a given theory to have a particular ontological commitment; or: how do we decide what the ontological commitments of a given theory are? This is the question that W.V. Quine tackles in his famous paper "On What There Is."[1] Quine's answer is that a theory is ontologically committted to those things which the bound variables of its sentences must be considered to range over for those sentences to be true.[2] Here there is an implicit assumption that the theory is expressed in a canonical

notation of quantifiers and bound variables, but this is no problem since presumably all theories worth having can be put into that form without any difficulty (i.e. into some notation of quantifiers and bound variables, not necessarily Quine's, which might be regarded as too restrictive for certain purposes). Another point is that for 'sentences' we should understand here 'unanalyzable sentences'. The theory may contain sentences that look as if they require the postulation of certain kinds of objects for their bound variables to range over, but which can in fact be analyzed into sentences for which this is not true at all, and, as Quine himself recognizes, we only want to accept the apparent commitments involved in an analyzable sentence if they are also involved in its analysans.

An example which Quine gives of the application of the criterion involves considering a theory which contains, or entails, the sentence 'Some species are cross-fertile'.[3] Assume that the sentence is unanalyzable. If the theory is expressed in the canonical form of quantified logic, the sentence will have the structure $(\exists x)$ $Sx \,\&\, Cx$. Now it is clear that if we accept the theory in question, then on Quine's criterion, we must suppose that there are things wich are species, (i.e. universals like 'Lion' 'Homo Sapiens' etc. not just the individual creatures that are instances of these universals), and things which

are cross-fertile, since such things are clearly needed for the variable 'x' in the above sentence to range over for the sentence to be true. But, Quine points out, the fact that the theory entails this sentence does not mean that it commits us to the universals 'specieshood' and 'cross-fertility'. The sentence does not contain bound variables which need to range over these latter things. Here, then, we have a clear example of how Quine's criterion can be used to distinguish between ontological commitments which a theory only seems to have and those which it really does have.[4]

One aspect of Quine's criterion that one might wish to criticize is its failure to take account of what might be called 'conditional' ontological commitments as well as 'categorical' ones. Such conditional commitments exist and are often very important. For example, modern chemical theory is not categorically committed to the existence of plutonium. This theory would still be true even if there were no plutonium, for it only tells us what properties plutonium will have if it exists. But it could be said to be conditionally committed to the existence of plutonium in the sense that certain possible facts concerning the properties of certain materials together with portions of the theory concerned with the recognition of substances from their properties would support the belief that

plutonium existed (support but not entail, since it would always be possible to interpret the results of the test for plutonium in some other way). What this seems to suggest is that we should regard Quine's criterion as a criterion of categorical ontological commitment only and that we need to add to it a criterion of conditional ontological commitment which might run as follows: a theory is conditionally committed to ϕ_s when some possible state of affairs would, if one accepted the theory, support a statement unanalyzable within the theory which (i) contained a bound variable that had to be regarded as including ϕ_s within its range and which (ii) would not be supported by that possible state of affairs alone.

Simplicity recommends a further modification to Quine's conception of the matter. This modification produces the following for the criterion of categorical commitment: a theory is categorically committed to ϕ_s when it entails '$(\exists x)\phi x$' as an unanalyzable sentence. And for conditional commitment we get: a theory is conditionally committed to ϕ_s when a possible state of affairs, taken with the theory itself, supports '$(\exists x)\phi x$' as an unanalyzable sentence and the latter is not supported by that possible state of affairs alone. These criteria seem to embody a considerable simplification without departing from the spirit of Quine's original.

I think that this Quinian notion does indeed capture part of what is involved in believing that there exists a certain category of objects. But I now wish to raise doubts about whether it captures all of what is involved. I wish to argue that what objects a theory is committed to depends not just on the content of the theory but also on the kind of semantics that is given for the expressions of the theory.

Physics, on the face of it, seems ontologically committed to such objects as neutrinos. But if we refuse to give as our semantic theory for the language of physics one which, intuitively speaking, makes the statements putatively referring to such objects genuinely refer to them -- for example a theory like that which Davidson and other writers have been developing from Tarski's theory of truth[5] -- then arguably the commitment to these objects is only apparent. (I shall henceforth refer to such theories as 'representational'). It will rightfully be asked at this point what kind of semantics we could use instead. One possibility is that we simply let the theory which uses this language be its own semantics -- that we recognize no semantics for the language of the theory over and above the theory itself. For example, to explain the semantic functioning of the word 'neutrino' in physics we would on this plan simply say what physics says (in general) about neutrinos.

(It is not absolutely accurate to say that, according to this method, we let the theory be its own semantics. For the theory will talk about (say) protons (or at least appear to talk about protons) while the semantics will talk about sentences in which words like 'proton' will occur. It would be more accurate to say that the semantics consists of a mentioning of every sentence that the theory itself uses. So when we say that the theory and its semantics are the same, this is only a rough way of talking). Now it is implausible to suppose that a semantic theory for the expressions of physics could be totally formalistic. There would have to be links of some sort with observation and experience. But it is reasonable to suppose that highly theoretical expressions need not be directly related to anything but other expressions and that it is only the observational expressions themselves that are linked to anything 'outside' the theory itself.[6] So with respect to such theoretical expressions the semantics can be said to be non-representational. In that case even if the theory either contains or entails the sentence '$(\exists x)\ \phi x$'

(where 'ϕ' is a theoretical predicate), it does not have to be regarded as ontologically committed to ϕs. For if this kind of semantics is acceptable for the theory, then one is simply using the expression 'ϕ' in accordance with certain rules governing what sentences one can assent to, given that one has assented (or would be prepared to assent) to certain other sentences. One is not referring to literally real ϕs.

If the above is true then the conditions for escaping an ontological commitment are not as strict as is sometimes thought. An ontological commitment is not the sort of thing that it is easy to be entrapped in against one's will. To avoid the commitment it is not necessary to drop the theory which ostensibly talks of the entities in question. It is true that one cannot accept a certain kind of semantic theory for the expressions one uses and avoid the commitment, and this certainly constitutes a restriction. But more often than not part of one's reason for choosing a particular semantic theory will be precisely in order to avoid, or to accept, a particular ontological commitment. Nor is it necessary, in order to avoid an ontological commitment, to carry out a 'reduction' which involves showing that we can say all that we want to say without using the offending expressions. All that one needs to do is to show that the sentences of one's theory are not to be

interpreted by means of a representational semantics. The only thing that one can commit oneself to in the mere use of a certain theory, independently of any considerations concerning the kind of semantics one would accept for it, is that one will affirm certain statements only when certain other statements are affirmed. How is it possible to read a genuine ontological commitment out of this? It might be objected that a person who uses physics without accepting an ontological commitment to neutrinos does not literally believe the theory he is using. He is merely behaving as if he thought physics were true. But I will not quarrel with this. If it is accepted, then it only means that my point will have to be put slightly differently -- namely, by saying that really believing in physics (and hence really accepting an ontological commitment to such things as elementary particles) differs from merely using it without really believing it (and hence not really accepting such a commitment) in the kind of semantics one is prepared to accept for the sentences of physics -- and therefore that ontological commitment still depends upon the kind of semantics that one is prepared to offer for a theory.

Let us return for the moment to the issue of ontological reduction briefly alluded to in the previous paragraph. It turns out in fact that there is a very simple way in which one can, in most cases, carry out a reduction

if one wishes to. In the first place one need never assert any of the statements that constitute the theory. For each such statement S, instead of asserting S one can assert the metatheoretical statement that S is a theorem of T (where T is the name of the theory in question). Thus one easily avoids any ontological commitment that might have been involved in asserting those statements themselves, since the offending expressions will now only occur within quotation marks. Secondly, whenever in using the theory one wishes to make a contingent assertion employing the concepts of the theory (for example, in the case of physics, that there is an electron at such-and-such coordinates), one can replace the sentence S in question by the sentence 'S is now affirmable according to T'. For the same reason as before, any apparent commitment to entities referred to in S drops out.

There is clearly, however, a limitation on the use of this technique: one cannot use it when the class of objects that one wishes to eliminate is one which includes sentences or theories. Thus even when there is no genuine ontological commitment, one cannot in general perform an elimination of the expressions which seem to carry it. But this should not worry us, for, as we have seen, the escape from the commitment depends upon something other than the eliminability of the relevant expressions. One should not conclude from this, however, that 'reduction programs' are

a waste of time. It is very interesting to know how simple and linguistically economical we can make a theory. It is simply that we should not regard such programs as having the ontological implications that they are usually thought to have.

There is a certain possible misconception one can have about the criterion of ontological commitment being suggested here. One might think that the reason why giving a 'formalist' theory of meaning lets us off the hook ontologically speaking is that although the language for which the theory of meaning is given talks about, say, electrons, the theory of meaning itself does not (it merely talks about the word 'electron' and the inferential relationships which sentences containing that word have to other sentences), so that there does indeed seem to be some sort of elimination of electrons themselves. But it is clear that to say this would be to commit again the very error that we have already tried to expose -- namely, that of thinking that just because a theory entails '$(\exists x)\phi x$' it is thereby committed in the literal sense to the existence of ϕs. Suppose indeed that a certain theory T entails '$(\exists x)\phi x$'. To know whether there is a genuine commitment to ϕs we must ask what kind of semantic explanation is appropriate for the use of 'ϕ' as it occurs in T. Suppose we hit on a certain non-representational explanation T', but T' itself entails '$(\exists x)\phi x$'. Again in order to see whether

there is a genuine commitment to ϕ_s involved, we must ask about the semantic functioning of 'φ', this time as it occurs in T'. If the explanation is representational, then there is indeed a genuine commitment to ϕ_s. But quite possibly we may feel that to explain the use of 'φ' in T', we need only re-apply the principles used in explaining its use in T -- i.e. non-representational principles. So there is still no genuine commitment to ϕ_s. The mere fact that T' entails '$(\exists x)\phi x$' does not mean that we have not escaped that commitment. What is crucial to the question of ontological commitment to ϕ_s is the representationality or non-representationality of the theory of meaning involved, not the latter's reference or lack of reference to ϕ_s.

Let us take a particular example of the situation described above. Suppose our theory of meaning for the language of physics is partly 'formalist' and partly 'non-formalist' in the following manner: 'theoretical statements' are not given an interpretation (their 'meaning' is simply their 'place in the system'), but 'observational statements' are, their interpretation being in terms of Quine's concept of 'stimulus-meaning'.[7] Does this type of theory of meaning relieve physics of its commitment to, say, elementary particles? The first thing to notice is that the concept of stimulus-meaning makes reference to human behaviour. Now it is true that the concepts in terms of which human behaviour

is described may not be the concepts of physics themselves. We do not have to think of the movements of human bodies as entailing (in the strictest sense) the movements of vast aggregates of atoms. We may use instead an ordinary pre-scientific concept of human behaviour. If we do, it will be relatively easy to concede that we are not committed to the elementary particles, because our reference to them has been 'discharged'; it does not occur in our semantic account of sentences in which the relevant expressions occur. But suppose instead that we do understand the behaviour in terms of the very physical concepts the nature of which it was the purpose of our theory of meaning to elucidate. Then it might be thought that we have not after all 'conjured away' the 'referents' of those concepts -- the elementary particles and so on. But it is clear that if the reasoning of the previous paragraph is sound, then that thought would be incorrect. And yet we may still feel uneasy. And this uneasiness is perhaps justified if it arises out of a doubt as to whether this is an adequate kind of semantic theory for the language of physics, but not if the doubt concerns whether or not the ontological commitment is discharged if it is adequate. Reasons will be given later for supporting that first doubt (it is clear to begin with -- and this is perhaps what causes the uneasiness referred to above -- that there is some kind of circularity involved in the explanation;

though merely to point this out is not, as we shall see, enough).

Another possible misunderstanding that one might fall into is that of thinking that we have brought to light two different 'kinds of reality' -- one associated with representational theories of meaning and the other associated with non-representational theories of meaning. This would perhaps do as a loose statement of our position, but it would be better to think of there being two quite different senses of the word 'real' which have very little in common, rather than two kinds of reality.[8] It is also somewhat misleading to say that we are arguing that we can avoid our apparent ontological commitments by treating the objects concerned as 'useful fictions'. For all that is needed in using a non-representational language is that one follow the rules of the language. Of course it may be psychologically helpful, when one uses certain expressions of the language, to imagine them as referring to entitites that have the same full-blooded kind of reality as is associated with representational language, but it is not essential. If one uses a language in which there is ineliminable quantification over, say, sets, then one gets out of attributing literal reality to sets (if one wants to get out of that) by refusing to give a representational semantics for the set-expressions of the language, not by substituting 'imagined reality' for

'genuine reality'.

The conclusion of our discussion may be expressed thus: a theory is categorically committed to the ('full-blooded') reality of ϕs where (i) it entails '$(\exists x)\phi x$' as an unanalyzable sentence and (ii) the use of 'ϕ' in the theory is to be explained representationally (I shall omit the definition of 'conditional commitment' which is an obvious extension of this). It might be asked just why we base the notion of ontological commitment so crucially on the notion of a representational semantics. In fact it is very difficult to find an argument for this. But hopefully the use of an argument will be conceded to be inappropriate. All we can do is present the criterion to the reader and simply ask him to judge whether or not it brings out what he had in mind when he thought about 'ultimate' or 'honest-to-God' reality. It is my belief that if he also bears in mind the accompanying comments that appear in this chapter, the judgement will be affirmative.[9]

In this chapter we have only considered what is involved in believing a certain category of objects to be truly real. We have not considered the question of what objects we should regard as real. My answer to this is experiences, and experiences alone. This answer is defended in the fourth chapter.

CHAPTER TWO: SKETCH OF A NON-REALIST APPROACH TO SEMANTICS

1. INTRODUCTORY REMARKS

In this chapter we shall describe and illustrate an approach to semantics which is non-representational with respect to all expressions other than experiential ones and which, in view of the findings of the previous chapter, is therefore non-realist with respect to everything except experiences. It should be pointed out that our concern at the moment is almost entirely expository. We leave the treatment of major criticisms and doubts to the last chapter -- the method must be fully understood before we can see which criticisms are to the point.

One of the crucial features of the theory which I advocate is that it is not 'reductionistic,' that is, it does not seek to provide a method of translating sentences apparently not about experiences into sentences that clearly are about experiences. This distinguishes it from the theory known as 'phenomenalism'. It is well known that phenomenalism is extraordinarily difficult to maintain particularly because of considerations having to do with its reductionistic character. I have no hope of being able to bolster it up. We shall see later how the theory which I advocate manages to cope better with some of the difficulties raised by phenomenalism.

However, there is one important feature which the present method does share with phenomenalism. This is its conception of the nature of experience. Phenomenalism conceives experience in terms of 'sense-data' -- the 'ideas' of Locke and Berkeley and the 'impressions' of Hume. In recent years these entities have become very unpopular. However, I believe that unlike the arguments against reductionism, the arguments against sense-data themselves can be countered. I shall try to show this in my last chapter. I am, then, advocating a form of anti-realism with respect to the 'external world' that takes sense-data as basic, though not in the sense of effecting a straightforward translation of non-sense-datum statements into sense-datum statements. We now proceed to an account of how anti-realism can be maintained without such translatability.

2. ACCEPTABILITY

In this work we confine ourselves to assertoric language i.e. language which, intuitively speaking, is used to indicate states of affairs. A key concept in our analysis of such language is that of the acceptability[1] of a sentence relative to a given set of experiences. This is a number that indicates the likelihood that the sentence in question is true given only that all the members of the set do indeed occur. It is via this notion that our semantic theory is given its pragmatic significance. For the appropriateness of uttering a certain sentence in the assertoric mode depends to a large extent on the acceptability of the sentence relative to some set of experiences - perhaps the set of all experiences the utterer could reasonably be expected to think of bringing to bear on the question of the sentence's truth. A further reason for making acceptability[2] basic is that, as we shall see, it can be used to define 'truth' (despite the fact that we actually invoked 'truth' above to give a brief intuitive explication of acceptability). And it is acceptability that enables us to be anti-realists with respect to non-experiential entities without resorting to phenomenalistic reductionism. For what we shall require semantic principles to do is not to tell us whether the occurrence of any given

set of experiences is a necessary and sufficent condition for a non-experiential sentence's truth, but rather what the acceptability of such a sentence is relative to that set of experiences. It seems clear that if the semantic principles governing non-experiential language achieve no more than that - or, strictly, as we shall see, do no more than combine with inductive and deductive principles to achieve that -- then such language is being treated no more representationally than it is in the phenomenalist's analysis.

We shall begin by stating more exactly what our notion of acceptability relative to a set of experiences is intended to represent. In general the acceptability of p relative to the set $\{x_i\}$ is the likelihood that p given all the characteristics of all the members of $\{x_i\}$.[3] Now suppose we wish to represent in this scheme the likelihood that p given that there exists an experience having the (possibly complex) property ψ. With a certain qualification to come later, this likelihood will be the acceptability of p relative to any set that has, as its sole member, a ψ (i.e. an experience with property ψ), except of course where there are no ψ s. If indeed there are none, it is the acceptability that p would have relative to any such set, if there were ψ s (by this means, we avoid having to include 'possible experiences' amongst the members of the

sets relative to which sentences have acceptabilities). Similarly, if there are experiences that have ψ and (distinct) experiences that have χ, the likelihood that p, given that there is an experience with ψ and a (distinct) experience with χ is, with the same qualification hinted at above, the acceptability of p relative to any set which has as its only two members a ψ and a χ; and if it is not the case that there exist distinct ψs and χs, it is the acceptability that p would have relative to any such set, if distinct ψs and χs did exist.

The 'further qualification' mentioned above is this: the likelihood that p given that there is a ψ (to take the simpler of the two examples) must be incapable of being affected by any other characteristics that the ψ might have (if this is not so, then this likelihoood cannot be represented in our scheme). For suppose this condition is not satisfied. Assume first there are ψs. The acceptability of p relative to any set whose one member is a ψ will depend on all the characteristics of that ψ. So only if p's likelihood given that there is a ψ could not be affected by other possible characteristics of the ψ will this acceptability represent the likelihoood that p given only that there is a ψ. Suppose alternatively that there are no ψs. The suggestion given above is that the likelihood that p given that there is a ψ is the acceptability that p would have

relative to any set containing only a Ψ if there were Ψs. But again this is only so if other possible characteristics of the Ψ could not affect the likelihood that p. For if this condition is not satisfied, there will be no one acceptability that p would have relative to any set containing only a Ψ if there were Ψs; in such a case one would get different acceptabilities for the different characteristics of the various possible Ψs.

Besides the notion of acceptability relative to a set of experiences there is also the important notion that I have chosen to refer to as the satisfaction of a sentence by a set of experiences. A sentence S is satisfied by a set of experiences E when an examination of all the members of E would render S certain.[4] Clearly, for S to be satisfied by E it must be experiential in content, and further it must be about some or all of the experiences that constitute E. In fact E satisfies S when S consists entirely in a correct description (whether partial or complete) of some or all of the members of E. If S does no more than attempt to describe members of E precisely then there is only one other possibility besides its being satisfied by E -- what I refer to as E's contrasatisfying it -- where a complete examination of all members of E would render it certain to be false (clearly S is contrasatisfied by E just in case $\sim S$ is satisfied by E). A semantic system will be so designed

that a sentence S can only have an acceptability relative to E if it both fails to be satisfied and fails to be contra-satisfied by E .

It might be wondered why, instead of invoking the notion of satisfaction, we do not instead talk of a sentence having maximum acceptability. The fact is that in our system there is no maximum acceptability. The full defence of this will have to wait until a little later to emerge. In the meantime we can at least note that while the conventional course might appear technically superior, the present one is perhaps better both from a philosophical and from a heuristic point of view. The situation in which a sentence is satisfied by an experience-set is qualitatively different from that in which it has an acceptability relative to that set. In our scheme, there are four distinct kinds of situation to be recognized. In the first two the sentence S consists entirely in a precise description of some or all of the members of the experience-set E. The first situation is that in which E satisfies S , the second that in which E contrasatisfies it. Here the question of S's acceptability relative to E -- of how likely S is on the basis of E -- does not arise (or more accurately we should adopt the Russellian approach and say that any sentence attributing a value to the acceptability of S relative to E ($A(S,E)$) is false because $A(S,E)$ does not exist). In the third

kind of situation S transcends a mere description of the content of E but E bears upon its truth to a greater or lesser degree -- here $A(S,E)$ does exist. Finally, in the fourth type of situation E is simply irrelevant to S, so as before $A(S,E)$ does not exist.

How in general is $A(S,E)$ determined? Here we must introduce the notion of an inferential factor of S relative to E. These are of two kinds: positive inferential factors (or 'positive factors' for short) and negative inferential factors (or 'negative factors'). A positive factor of S relative to E is a prima facie reason for believing that S is true on the basis of E. A negative factor of S relative to E is a prima facie reason for believing that S is false on the basis of E. Each inferential factor of S will consist of a set of premisses and the conclusion, S itself (or $\sim S$ in the case of a negative factor). The premisses are sets of analytically equivalent sentences. (They may thus be viewed almost as 'propositions'.) $A(S,E)$ will be determined by the strengths of the inferential connections involved in the various positive and negative factors that S has relative to E and by certain properties of the members of their premisses. Deductive and inductive rules will tell us that when there exist sentences $S_1, S_2 \ldots S_n$ with certain properties and a further sentence S bears a certain relation to these sentences, S will have an inferential

factor relative to E of a certain strength with $\alpha_1, \alpha_2 \ldots \alpha_n$ as premisses, where each α_i is the set of all sentences

and negative factors of which their negations are conclusions. It should be noted that once S and E are fixed and the rules of the system are fixed, $A(S,E)$ is fixed too. S only has an acceptability relative to E in virtue of one or more applications of rules of the system. Acceptability cannot transcend the operation of these rules; in effect, the rules define what acceptability is.

It might be wondered why the premisses of inferential factors are not sentences but sets of analytically equivalent sentences. The answer to this is very simple: we do not want to count separately an inference from, for example, p to q and an inference from $p \vee p$ to q. On our method both of these inferences will be subsumed under the inferential factor whose premiss is the equivalence-class containing both p and $p \vee p$ and whose conclusion is q. As we shall see later, this problem of preventing duplication of inference-patterns will necessitate other more complex stipulations at certain points.

We are now in a somewhat better position to defend the absence of an upper bound to our acceptabilities. Suppose we have a sentence S which is supported by a number of inductive positive factors. Suppose also that we keep on discovering further inductive positive factors for S. For the acceptability of S to have an upper bound, either the new factors must become redundant beyond a certain point

or the extent to which the addition of a new factor contributes to the acceptability of S must be constantly diminishing. But this seems highly artificial. At what point should the redundancy begin -- or, if there is a constant diminution in the weight of the successive contributions, what will the rate of diminution be? The answer to either question can only be arbitrary, it seems. There would appear to be a confusion here between the probabilistic facts themselves and the use we make of them. We may decide for practical reasons to ignore further evidence when the evidence we already have is very good. But such a tendency should not be incorporated into the probabilistic system itself. Of course, the rejection of a maximum acceptability does have the unfortunate consequence of making it impossible to identify acceptabilities with mathematical probabilities.

A system for determining the acceptabilities of sentences relative to sets of experiences will have at least the following ingredients:

(i) A representational theory of the meaning of certain precise experientialexpressions. This might take the form of a Tarski-Davidson-style truth-conditions theory. A set of experiences E will satisfy a sentence S just in case the obtaining of the truth-condition associated with S by the representational theory is rendered certain by the totality

of characteristics of all the members of E , and will contrasatisfy it just in case its non-obtaining is rendered certain by the same.

(ii) A series of semantic principles governing both logical and non-logical expressions. These may take the form of straightforward axioms or axiom- schemas or they may be rules attributing positive and negative factors to sentences when there exist other sentences supporting or undermining them.

(iii) Inductive rules. These do the same as the above, but on the basis of inductive rather than semantic connections.

(iv) A definition of 'deductive consequence'. This will provide the basis for the notion of a 'theorem' (a deductive consequence of some of the axioms, whether logical or semantic) and for the notion of 'analytically equivalent' (P and Q are analytically equivalent when $\ulcorner P \equiv Q \urcorner$ is a theorem of the system).

(v) A Rule of Entailment. This is the deductive rule of the system whose approximate function has already been described on page 23. A more accurate characterization would run as follows: the rule states the conditions under which an experience-set E gives support to a sentence S in virtue of S's being a deductive consequence of some set α of sentences not containing S. There are two cases:

that in which the entailment of S by α makes it the case that S is satisfied by E and that in which the entailment creates a positive factor for S relative to E whose premisses are all the sets of analytically equivalent sentences containing a member of α . In the latter case the rule will also tell us precisely what properties of the members of α are relevant to (partly) determining $A(S,E)$, and will also indicate the fact that the strength of the actual inferential connection is the maximum possible. Obviously the rule appeals to the notion of 'deductive consequence' grounded in (iv) above.

(vi) A method of determining $A(S,E)$ on the basis of the inferential connections involved in the various inferential factors that S has relative to E and of the relevant properties of their premiss-members.

(vii) Definitions of all expressions that can be regarded as definable in terms of other expressions. These will consist of all expressions not dealt with in any other part of the theory or more specifically in the semantic constituents (i) and (ii). Clearly many of our notions must be extended in obvious ways to take account of defined expressions e.g. if the definiens of S is satisfied by E, so is S .

An important step in the articulation of the present method would be to say what constituent (vi) consisted in,

that is to say, how $A(S,E)$ is determined on the basis of S's inferential factors. This, however, seems to involve formidable difficulties which I shall not attempt to resolve. But what I will do is go into a little more detail about the Rule of Entailment.

All applications of the Rule of Entailment involve the following kind of situation: a set of sentences α entails a further sentence S not contained in α . Now there are two ways of further particularizing this situation to each of which corresponds a part of the Rule of Entailment (it should be noted that these two particularizations are not exhaustive and so not <u>every</u> case of a set of sentences α entailing a sentence S not contained in α is suitable for the application of the Rule of Entailment). The first is as follows:

(A) Suppose at least one member of α is satisfied by E and every member of α that is not satisfied by E is a theorem of the system.[5]

In this case what the Rule of Entailment tells us is very straightforward. It is simply that S itself is satisfied by E .

To describe situation (B) we need a definition. <u>The Acceptability of S relative to E without inferential factors $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n$ and without sentences $T_1, T_2 \ldots T_n$</u> (which may be written '$A(S,E)/\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n, T_1, T_2 \ldots T_n$' is the

number obtained when the acceptability of S relative to E is calculated neglecting the inferential factors $\xi_1, \xi_2 \ldots \xi_\lambda$ that S has relative to E and neglecting any inferential factor that S may have relative to E whose existence is established by a process that makes reference to any of $T_1, T_2 \ldots T_\lambda$ or their negations or any sentence logically or semantically entailing them or their negations.

Situation (B) is the following:

S is not satisfied by E ; at least one member C of α is such that $A(C,E)/S$ exists and is not entirely based on negative factors; every member of α of which this is <u>not</u> the case is either satisfied by E or is a theorem of the system; and finally no proper subset of α entails S .

This is the kind of situation in which we get a positive factor for S in virtue of its entailment by α (though as we shall see in a moment there are further conditions to be satisfied). There are two things to explain. First: why the stipulation that no proper subset of α entail S ? This is straightforward. We can add to any set of sentences that entails S any quite irrelevant sentence, and the resulting set will also entail S -- and clearly we do not want such a sentence to affect the acceptability of S . The requirement ensures that α is <u>minimal</u> in the sense that every member of it is essential for producing S as a deductive consequence. Second: why

do we require that at least one member C of α be such that $A(C/E)/S$ exists rather than simply $A(C,E)$? This is to prevent S supporting or undermining itself. In general, any rule which generates an inferential factor on the basis of sentences that have acceptabilities must require that these sentences have acceptabilities <u>without</u> the proposed conclusion of the factor.

We must now introduce the further conditions that have to be imposed on α and S for there to exist a factor of the kind in question. These conditions are intended to prevent duplications of essentially similar inference-patterns. Suppose α is $\{p, p \supset r\}$, S is r, r is neither a theorem nor satisfied by E, and both p and $p \supset r$ have acceptabilities relative to E without r neither of which is based entirely on negative factors. Then the conditions of situation (B) are clearly satisfied. But suppose the acceptability of $p \supset r$ without r is entirely based on its entailment by the set $\{p \supset q, q \supset r\}$ both members of which have acceptabilities relative to E without r that are not based entirely on negative factors. Then the union of the latter set with $\{p\} \cup (\{p, p \supset q, q \supset r\})$ can itself be taken as a set that satisfies, in relation to r, the conditions being imposed on α and S in the description of situation (B), but we do not want an inferential factor for r based on its entailment by $\{p, p \supset r\}$ <u>and</u> another based on its entailment

by $\{p, p \supset q, q \supset r\}$. We refer to this kind of situation as 'Total Duplication Type A in relation to $\{p, p \supset r\}$ and r.' It necessitates rejecting the possible factor for r based on its entailment by $\{p, p \supset r\}$. In the case we have described, the inferential factor will be attached to $\{p, p \supset q, q \supset r\}$, <u>provided</u> that no total duplication (either of type A or of type B or C to be described below) occurs in relation to <u>this</u> set and r. The method has the effect, roughly speaking, of bringing us back to the most complex set from which r can be derived.

Now suppose we take α to be $\{p, p \supset q, q \supset r\}$ and S to be r and we assume that p and $p \supset q$ both have acceptabilities relative to E without r that are not entirely based on negative factors, that $q \supset r$ is satisfied by E, and that r is not a theorem and is not satisfied by E. Then α and S satisfy the conditions of situation (B). But suppose $p \supset r$ is satisfied by E. Then we want the factor for r to be based on its entailment by $\{p, p \supset r\}$ not on its entailment by $\{p, p \supset q, q \supset r\}$. The description of this case would be 'Total Duplication Type B in relation to $\{p, p \supset q, q \supset r\}$ and r.' It necessitates rejecting the possible factor for r based on its entailment by $\{p, p \supset q, q \supset r\}$. Unlike in the former case, it is the simpler, not the more complex, set that we prefer the factor to be associated with.

Now consider the following situation: the set

$\{p \& q, p \supset r\}$ if taken as α, satisfies the conditions of situation (B) with r taken as S and therefore looks as if it might give rise to a factor for r. Suppose also that the grounding for $p \& q$ consists in an acceptability relative to E without r and that this rests on separate evidence for p and for q independent of r. Clearly $A(p,E)/r$ will be greater than $A(p \& q, E)/r$. We would therefore be making our case for r unnecessarily weak if we were to use the set $\{p \& q, p \supset r\}$ instead of the set $\{p, p \supset r\}$. We therefore choose the latter rather than the former. This kind of situation would be referred to as 'Total Duplication Type C in relation to $\{p \& q, p \supset r\}$ and r.'

Based on these considerations the further condition which the Rule of Entailment imposes on α and S for there to exist a positive factor for S based on its entailment by α (whose premisses are all the sets of analytically equivalent sentences containing a member of α) is that none of the following three situations obtain:

Total Duplication Type A in relation to α and S[6] occurs when there is a set of sentences β and a sentence P[7] such that:

(a) $P \in \alpha$

(b) β entails P

(c) $\beta \neq \{P\}$

(d) $\beta \cup (\alpha \cup \{P\}')$ satisfies in relation to S all

the conditions so far imposed on α in relation to S.

(e) $A(P,E)/S$ exists.

(f) $A(P,E)/S$ is entirely based on the entailment of P by β.

Total Duplication Type B in relation to α and S [6] occurs when there is a set of sentences γ and a set of sentences λ [8] such that:

(a) γ is a proper subset of α

(b) $\lambda \cup (\alpha \cap \gamma')$ satisfies in relation to S all the conditions so far imposed on α in relation to S.

(c) Every member of λ is either satisfied by E or is a theorem.

(d) Not every member of γ is either satisfied by E or is a theorem.

Total Duplication Type C in relation to α and S [6] occurs when there are sentences R_1 and R_2 [9] such that:

(a) $R_1 \in \alpha$

(b) R_1 entails R_2 but is not entailed by it.

(c) $\{R_2\} \cup (\alpha \cap \{R_1\}')$ satisfies in relation to S the conditions so far imposed on α in relation to S.

(d) $A(R_1,E)/S$ and $A(R_2,E)/S$ both exist and the former is smaller than the latter.

A finalaspect of the Rule of Entailment that we shall consider is the question of what property of a premiss-member M of an inferential factor of S relative to E is

relevant to the determination of $A(s,E)$ in the case where that premiss-member is neither a theorem nor satisfied by E . One might think it would always be simply $A(M,E)/s$ This is unfortunately not true. Let us go back to the situation described on p. 30, but modify it in one respect: let us suppose that $A(p \supset r, E)/r$ is based not <u>entirely</u> on the entailment of $p \supset r$ by $\{p \supset q, q \supset r\}$ but only partly -- there is other evidence for $p \supset r$ as well. Suppose also that $\{p, p \supset q, q \supset r\}$ enjoys no kind of total duplication in relation to r and so there is a positive factor for r based on its entailment by that set. In this case, provided that there is nothing that would give rise to any kind of total duplication in relation to $\{p, p \supset r\}$ and r , we have what is described below as 'Partial Duplication' in relation to $\{p, p \supset r\}$ and r , and although we want there to be a positive factor for r based on $\{p, p \supset r\}$ we do not want to take $A(p \supset r, E)/r$ as the relevant property of $p \supset r$ to be considered in assessing $A(r,E)$ since that would include the contribution made to $p \supset r$ by $\{p \supset q, q \supset r\}$ which would be duplicating the effect of $\{p, p \supset q, q \supset r\}$. Rather our representation of the 'strength' of $p \supset r$ must (for the purpose of this assessment) neglect the factor that it owes to its entailment by $\{p \supset q, q \supset r\}$. This situation can also be used to illustrate what I have chosen to call 'Internal Circularity'. If $\{p, p \supset q, q \supset r\}$ is not totally duplicated in relation

to r , and if $p \supset q$ gains some kind of support from $p \supset r$, then we have a case of Internal Circularity in relation to $\{p, p \supset q, q \supset r\}$ and r . The appropriate response in this case is to neglect the support that $p \supset q$ gets from $p \supset r$ in our representation of the strength of $p \supset q$ as a premiss-member of the positive factor for r based on r's entailment by $\{p, p \supset q, q \supset r\}$, otherwise we are, as it were, supporting r by a process in which $p \supset r$ supports itself (in detail the situation would be this: $p \supset r$ supports, say, $p \supset q$; $p \supset q$ combines with $q \supset r$ to support $p \supset r$; $p \supset r$ combines with p to support r .

We must now formally define Partial Duplication and Internal Circularity.

Partial Duplication in relation to α and S [6] occurs when total duplication does not occur in relation to α and S , but there are sets of sentences ω and β and a sentence P [10] such that:

(a) $P \in \alpha$

(b) β is a subset of ω

(c) $A(P, E)/S$ exists and is based partly, but not entirely on the entailment of P by β .

(d) There is a positive factor for S based on its entailment by ω .

(e) ω entails every member of α .

Internal Circularity in relation to α and S [6] occurs

when total duplication does not occur in relation to α and S, but there is a set of sentences γ and a sentence Q [11] satisfying the following conditions:

(a) γ is a proper subset of α

(b) γ entails Q

(c) $\gamma \neq \{Q\}$

(d) $\{Q\} \cup (\alpha \cap \gamma')$ entails S.

(e) A member of γ is supported in some way by Q.

We must now state which property of a premiss-member of a positive factor relative to E based on α's entailment of S is relevant to the determination of $A(s,E)$.

If partial duplication obtains in relation to α and S, the relevant property of the premiss-member M is $A(M,E)/\xi_1, \xi_2 \ldots \xi_n, S$ where $\xi_1, \xi_2 \ldots \xi_n$ are all the inferential factors based on the entailment of a P by a corresponding β (P and β as defined on p. 35 above).

If internal circularity obtains in relation to α and S the relevant property of M is $A(M,E)/\Omega_1, \Omega_2 \ldots \Omega_n, S$ where $\Omega_1, \Omega_2 \ldots \Omega_n$ are all the positive factors pertaining to the support of a member of γ by a corresponding Q (γ and Q are defined on p. 36 above).

If partial duplication <u>and</u> internal circularity occur in relation to α and S, the relevant property of M is obviously $A(M,E)/\xi_1, \xi_2 \ldots \xi_n, \Omega_1, \Omega_2 \ldots \Omega_n, S$ with the ξ_is and Ω_is as defined above.

3. SEMANTIC PRINCIPLES

Up to now we have only been concerned with the bare structure of the theory. We now fill in that structure to some extent by indicating the sort of principles that would be invoked to explicate the use of non-experiential expressions. Let us reiterate that the method differs from reductionistic phenomenalism. We do not seek rules that will enable us to state necessary and sufficient experiential conditions for a non-experiential sentence's being true (such conditions do not, in general, even exist[12] -- more will be said on this later). Instead our rules are intended to determine the acceptability of any non-experiential sentence relative to any given set of experiences whenever such an acceptability exists. However, we shall not attempt to develop an exact and detailed system of rules. That would be too complex a task for the present work. Rather we will try to make suggestions as to the kinds of rules that might be included, focussing attention on particular kinds of discourse. Hopefully this will be sufficient at least to make it plausible to suppose that the task could be carried out in detail.

Color-Predicates

We begin with a particular class of expressions

which are relatively closely connected with experience -- color adjectives as these are applied to physical bodies. If we were to try to give an experiential definition of 'object x is blue at time t' our first attempt might be (assuming that we do indeed want to understand experiences in terms of sense-data) that an object is blue at t just in case any visual sense-datum of the object occurring at t would be blue. But clearly this will not do as it stands. An object can be blue without looking blue and can look blue without being blue. A phenomenalist might try to define 'x is blue at t' as 'for any observer y if standard conditions for color-vision obtain in relation to y and x at t, y will have a blue sense-datum at t.' Obviously the phenomenalist could only use this as a starting-point, for the definiens is not yet in purely experiential language. In particular, a purely experiential definition of 'standard conditions for color-vision obtain in relation to y and x at t' is required. But such a definition cannot be given, if only because there is no pre-set limit to the types of non-standard conditions that there can be. Things can go wrong in ways that are completely unforeseeable in principle[13] (this is clearly related to the comment made earlier about the absence of sets of necessary and sufficient experiential conditions for the truth of at least some non-experiential statements). The

method of dealing with color-words which is described below allows for this fact.

I have not been able to develop a formally precise axiom for color-words in the use we are concerned with. However, there follows an informally stated condition for an object to be blue at some given time, which could be used as the basis for such a formal axiom.

> An object x is blue at t just in case there is a set of conditions, each of which is relatively rarely fulfilled such that, if the non-fulfillment of all of them is designated as '**S**':
> (a) **S** , in conjunction with the statement that there exists at t a visual sense-datum of x, causally implies that that sense-datum is blue.
> (b) If **S** causally implies the occurrence of some change in the state of x occurring at or prior to t which will or would cause any visual sense-datum of x existing at t to be blue, the truth of (a) above does not depend on this fact.
> (c) The existence at t of a blue sense-datum of x causally implies **S** .

A comment should perhaps be made about the application of the word 'blue' here to both objects and sense-data. Obviously 'blue' differs in its meaning in the two applications ('blue' in its application to

sense-data, being an experiential expression, is explicated in the representational part of the theory; 'blue' in its application to objects is explicated by the axiom itself). In our informal statement there is no ambiguity, since it is always made clear whether 'blue' is being applied to an object or to a sense-datum. In a formal approach one of two possible procedures could be adopted: different styles of variables for objects and sense-data could be used, and the same predicate-symbol for 'blue' used in connection with both, or alternatively the same style of variable could be used, but with two distinct 'blue'-predicates.

Why clause (b)? This is explained by the fact that we want S to represent the condition for an object to appear the color that it is, and so we want to rule out the possibility that S makes an object look blue by actually <u>changing</u> its color. This is what clause (b) does.

Much doubt about the axiom may be engendered by its appeal to the concept of 'causal implication'. By 'p causally implies q' I mean that p materially implies q in virtue of some kind of causal connection between the two. But if a causal implication is conjoined to the negation of its antecedent the result is a 'subjunctive' or 'counterfactual' conditional, a type of statement concerning which there have arisen many problems. Yet

whatever the technicalities of the matter, the broad lines of an acceptable analysis of causal implications, and therefore also of counterfactual conditionals, seem evident enough. We know the kinds of situation that justify the assertion of a causal implication, and it is only a matter of achieving the (technically, but not philosophically, difficult) task of stating them in a simple and consistent systematic way. Such a systematization would include rules that would, for example, give a positive factor to 'If something is F then (causally) it is G' when high acceptabilities characterize both '$(x)\ Fx \supset Gx$' and also many (varied) positive instances of '$Fx \,\&\, Gx$.' However, this causal implication can also receive indirect support in cases where there are few known instances of '$Fx \,\&\, Gx$' or none at all and our rules would have to take account of this too. In addition, we would need to introduce further subtleties to explicate another concept which is appealed to in our axiom -- that of one event's actually causing another. However, there is no reason to suppose that these things could not be done.[14] Their being done would probably represent an important step forward in the explication of many predicates applicable to objects besides color-predicates. For many, if not all, such predicates seem to represent capacities of some kind and thus to be amenable to analysis in terms of causal implications (spatiotemporal predicates seem to be

particularly difficult expressions to pin down, but some of them at least could perhaps be taken as representing the capacity to present certain visual perspectives).

But let us now return to our axiom. How should we understand the term 'condition' as it occurs there? One possibility is that we take it to mean a sentence, but this would not be a sentence in the sense of a concrete utterance or written mark. Rather it would have to be a possible mark or utterance, (and clearly this reference to the possible would ultimately have to be cashed out in terms of the actual, just as the use of causal implications has to be) perhaps even one belonging to a language other than the one which the axiom is being used to explicate. For the set of non-standard conditions must, as we have already remarked, constantly be kept open -- capable of receiving further additions.

An example of the use of the axiom would be to show that if an object appears blue at t, then the safest assumption in the absence of further evidence (including evidence about the conditions of observation) is that it is blue at t. This can be seen from the following reasoning: if it is not blue at t (but does appear blue at t - and therefore does exist at t), then it must be some other color at t (it is true that we talk of colorless physical objects -- particularly liquids. But 'color' in the context of the axiom can be taken to include absence of 'color' in the normal sense). But this other color

(call it 'C') will also be governed by an axiom which tells us that if an object is C at t, then for some set of fairly rare conditions, the non-fulfillment of every member of the set, in conjunction with the statement that there exists at t a visual sense-datum of the object, causally implies that that sense-datum is C. But in that case the existence at t of a sense-datum of the object that is blue (i.e. not C) causally implies, by contraposition, the obtaining of one of the non-standard conditions which, however, would be a fairly rare occurrence. Hence it is more reasonable to suppose that the object is blue at t. In short, it is more likely, ceteris paribus, that an object is as it appears than that it is other than it appears. And if we knew also that one of the more common non-standard conditions for color-vision (i.e. common compared with other non-standard conditions for color-vision, not common overall) did not obtain, this would further strengthen our case for saying that none of them obtained, and therefore that the object was not C and hence (because this holds for any C other than blueness) that it was blue at t (since, to repeat, the object's being C at t and yet appearing blue at t would imply the obtaining of one of the non-standard conditions).

Identity-conditions

What we are confronted with in our perceptions are momentary (or at least rather fragmented) episodes. From these fragments we construct continuous objects. The manner in which the fragments are 'joined together' determines the kind of object we get (four-dimensional 'world-line' or ordinary common-sense three-dimensional object, for example). This in turn can be expressed by identity-conditions for objects. Thus an important role is played by identity-conditions in explicating the concept of a given type of object. All this follows Quine's paper "Identity, Ostension and Hypostasis"[15] and, to some extent, N.L. Wilson's "Space, Time and Individuals."[16] But what kinds of entities are the 'momentary episodes' referred to above? I think it would be a mistake to suppose that there _are_ such things unless they are identified with sense-data. Quine's conception of the matter, which appears to invoke momentary _physical_ entities, seems to be an implausible attempt to gain some of the advantages of sense-datum theory without embracing the theory itself. If we can accept sense-data at all (though I admit that this is something yet to be defended), _they_ can serve as the momentary episodes and we have no need of momentary physical entities. However, it should be noted that in practice we may have no _explicit_ mention of any process of compounding sense-data. If, for example, we say that physical

objects are identical just in case they are spatiotemporally continuous, the consequent of this biconditional asserts the existence of a continuous chain of physical objects. It is as evidence for the existence of objects at particular positions at particular times that sense-data become involved. Nevertheless, the overall effect is one of a synthesis of sense-data. Of course the identity-conditions that we actually finish up giving for physical objects would have to be more complex than a straightforward analysis in terms of spatiotemporal continuity. This would have to be worked on.

Natural Kind Words

Natural kind words appear to constitute a problem for an approach which tries to bind meanings firmly to the appearances of objects. Let us assume that we can indeed explicate all physical properties experientially. It might seem natural to define 'gold' say, in terms of a set of properties. But which properties should we choose? The totality of properties of gold which can be recognized by casual non-expert observation -- such properties as yellowness and hardness -- can be possessed by things which are not gold. On the other hand, the properties that experts use to distinguish gold from other substances seem inappropriate as constituents of the definition of gold just because they are known only to experts. This makes it seem as if one can do no more than say that 'gold' represents gold, i.e. adopt a

representational and therefore a realist approach. However, the objection against using the expert's definition can perhaps be countered. It is not too implausible to say that laymen do not know the 'true' meaning of the word 'gold'. It should be noted that not knowing the correct definition of a word is not incompatible with being able to use it competently in most situations. For such generally competent usage it is sufficient to be acquainted with the more obvious (and less reliable) properties of gold. But fully competent usage comes only with a knowledge of the expert's definition and so the latter has the right to be regarded as the 'correct' one. There is here, as Putnam has pointed out, 'division of linguistic labor.'[17] Thus once explications of the requisite physical properties have been achieved, the natural kind words can be defined in terms of these properties.

4. OTHER ASPECTS OF A SEMANTIC THEORY[18]

A brief comment should be made on a certain simplification that can be used sometimes in the calculation of acceptabilities. Suppose we are calculating the acceptability of the statement that one particular object *O* is blue at t relative to the set of all experiences of some person up to t. Suppose this set includes a blue sense-datum occurring at t. If good grounds can be given for thinking that this sense-datum is actually a sense-datum of *O*, then, if it is also thought likely that standard conditions obtain and in the absence of any further relevant evidence, a fairly high acceptability will be attributable to '*O* is blue at t'. But what kind of grounds might one have for thinking that the sense-datum is one of *O*? In fact, recognition of the sense-datum as being a sense-datum of *O*, recognition of an immediate, intuitive kind, may be quite adequate. This can be justified inductively if the observer's experiences up to t (i.e. the set relative to which the acceptability of '*O* is blue at t' is in question) include many instances of this feeling of recognition in relation to this type of statement, and among them, although there are many cases where the statement has been 'well-tested', there are few or no cases where its negation has a high acceptability relative to the entire set.[19] Thus we can make use of 'intuition' in the calculation

of acceptabilities, provided this intuition is legitimized by evidence of successful use in the past.

One doubt which may have occurred to the reader in regard to our method concerns the use of a concept in explicating another concept where the second does not seem to presuppose the first in the sense that it would be possible to grasp the second without having any grasp of the first. Thus, for example, our explication of 'blue' seems to require that any language containing the word 'blue' also contain the word 'cause'. This objection brings to light the fact that the use of an expression within the semantic theory does not necessarily mean that that expression occurs within the language which the theory is being used to analyze. There are therefore two kinds of sentences that may be assessed for acceptability: those which are well-formed within the language itself and those which, as it were, belong only to the theory, and whose acceptability is only important in determining the acceptabilities of sentences that <u>do</u> belong to the language as well as to the theory. Thus the language must have a vocabulary which is a subset of the entire list of expressions occurring in the theory. This means that the same semantic theory can sometimes correspond to several different languages. It also perhaps allows one to talk of concepts 'implicitly recognized' by a language that are not represented by its 'explicit' vocabulary.

CHAPTER THREE: TRUTH

1. TWO KINDS OF TRUTH

There is more than one way in which the word 'true' can be used. One important distinction is that between the use of 'true' as part of an operator which is concatenated with sentences (as in 'It is true that snow is white') and its use as a predicate which is concatenated with the names of sentences, or with variables ranging over sentences (as in 'The sentence "Snow is white" is true' or in 'Some sentences are true'). In this chapter we shall be exclusively concerned with the latter use of the word.

Even here, as we shall argue, there is at least one major philosophical distinction to be made. For consider a predicate 'ϕ' defined by the following principle:

> A sentence S is ϕ just in case there is a P such that P and such that S expresses P.[1]

It seems impossible to deny that if 'ϕ' is defined in this way it must constitute a truth-predicate (we shall henceforth refer to the truth that it represents as Redundancy Truth). But an argument can be produced for saying that it is not the only kind of truth to which we should pay attention. The argument runs as follows: consider a sentence of the form $\ulcorner p \vee q \urcorner$ where p and q are atomic sentences and both concern matters remote from direct experience, such as the position of some body in a distant part of the universe.

Imagine nevertheless that on the basis of some generalization we do have indirect evidence for $\ulcorner p \vee q \urcorner$ of a type which is not evidence for either of the disjuncts themselves (that is to say, we have some reason for believing $\ulcorner p \vee q \urcorner$. but we have no idea which of the disjuncts to believe). On this basis we might guardedly assert $\ulcorner p \vee q \urcorner$, and insofar as that is the case, we could also, by definition, guardedly assert that $\ulcorner p \vee q \urcorner$ had redundancy truth. But consider the following possibility: it turns out that relative to the totality of all experiences (i.e. the set that consists of every experience - past, present and future - of any experiencer, a set which we shall henceforth refer to as U , neither p nor q have a very high acceptability. This is a possibility which our indirect evidence for $\ulcorner p \vee q \urcorner$ does not give us any reason for disbelieving, particularly in view of the relative remoteness from experience of the subject-matter of p and q . Now it seems clear that if we look at the matter with an anti-realist attitude, then in at least one sense of 'truth', the possibility that we are talking about is incompatible with the truth of $\ulcorner p \vee q \urcorner$. On the assumption that U contains everything that there is, there cannot be anything in virtue of which $\ulcorner p \vee q \urcorner$ could be said to be true, if that possibility obtains. So there seems to be at least one sense of truth such that our evidence for $\ulcorner p \vee q \urcorner$ does not give us any reason for

disbelieving a possible state of affairs that is incompatible with the truth of $\ulcorner p \vee q \urcorner$ in that sense - and hence for believing that $\ulcorner p \vee q \urcorner$ has that kind of truth, but another sense of truth (redundancy truth) according to which our evidence for $\ulcorner p \vee q \urcorner$ does give us evidence for the truth of $\ulcorner p \vee q \urcorner$. There would therefore appear to be at least two extensionally different varieties of truth.

I shall refer to the new kind of truth as empirical truth. For a sentence to have empirical truth it would appear necessary that it either have a high acceptability relative to U, or be satisfied by U. But as the example just given indicates, this is not always sufficient. In the case of a disjunction, it must also be the case that at least one of the disjuncts has a high acceptability relative to U or is satisfied by U (in other words it is not sufficient that U furnish a large amount of merely indirect evidence for the disjunction that is not evidence for either of the individual disjuncts). This will be discussed more fully when we attempt a formal definition of empirical truth. What we need to observe at the moment is that not all classical logical theorems can be permitted to be empirically true. For example, a theorem of the form $\ulcorner p \vee \sim p \urcorner$ with atomic p will not be empirically true if neither of its disjuncts have a high acceptability relative to U or are satisfied by U - that is, if p is

indeterminate.[2] Does this mean that we have to abandon classical logic? I would argue not. Notice that every sentence of the form '$p \vee \sim p$' will have redundancy truth. We can do our semantics for classical logic in terms of redundancy truth rather than empirical truth, and thus still have a sound metatheory for classical logic. However, the retainability of classical logic actually needs more detailed defence, which it will receive presently.

How does the set of empirical truths relate to the set of redundancy truths? What would a Venn diagram representing these two sets look like? In fact it is not possible to answer this question since redundancy truth is really a disguised indexical or 'token-reflexive' notion. It is not an objective state that sentences either succeed or fail in getting into. It would almost be correct to say that 'S has redundancy truth' is equivalent to 'S is either analytic or the evidence available to me now favours S'. The only problem with this is that in the case of a synthetic S if one asserted on one occasion that S had redundancy truth and on another that it did not, then according to this construal one would not be contradicting oneself or changing one's mind. There seems to be another element involved - namely one of personal commitment to S.[3] In any case, it does seem wrong to say that redundancy truth represents any sort of objective state, in the same way as

empirical truth does (of course to say that there is no objective state involved is not to say that redundancy truth is a purely subjective matter. There are rules determining what can and cannot be asserted in various situations - so that if for example I claim that some sentence of the form $\ulcorner p \vee \sim p \urcorner$ is a redundancy truth and you deny it, my claim takes precedence over yours because I have applied the rules correctly whereas you cannot have done so. The point is merely that we should not think of the rules as determining any universally agreed state for sentences to be in). It is therefore not possible to compare the set of empirical truths with the set of redundancy truths. However, it is possible to say that we would like the set of sentences that we would be disposed to assert to have redundancy truth to be the set of empirical truths plus the set of theorems. So as far as non-theorems are concerned, empirical truth is the ideal end.

2. CAN CLASSICAL LOGIC BE RETAINED?

In this section I wish to examine an argument for maintaining that our principles do not permit us to keep classical logic. The argument rests upon the fact that there may be sentences that have neither empirical truth, nor empirical falsity (i.e. neither they nor their negations have empirical truth). It tries to show that Bivalance follows from the Law of Excluded Middle, and hence that any language conforming to classical logic can only admit a truth-predicate that satisfies bivalence. Since empirical truth does not satisfy bivalence, this would, if it were valid, require us to abandon classical logic if we wanted to hang onto empirical truth. Here is how the argument proceeds: ⌜S is true⌝ follows from S and hence so does ⌜S is true or ~S is true⌝; ⌜~S is true⌝ follows from ~S and hence so does ⌜S is true or ~S is true⌝; thus by the constructive dilemma principle, ⌜S is true or ~S is true⌝ follows from ⌜S ∨ ~S⌝. But the former is equivalent to the assertion that every statement is either true or false. So if ⌜S ∨ ~S⌝ were accepted as a theorem the possibility of indeterminacy would have to be denied. The flaw in this argument is that as far as empirical truth is concerned, ⌜S is true⌝ does not follow from S .[4] For a start, if S is analytic, the principle obviously fails, since analytic sentences are not empirically true. Moreover, the principle often fails to apply even to truth-functionally

complex synthetic sentences, as can be seen from the example on pp. 49-51, where evidence that supports $\ulcorner p \vee q \urcorner$ does not support the empirical truth of $\ulcorner p \vee q \urcorner$. However, the argument is easily modified to take account of these facts. Let us define $\ulcorner P(S) \urcorner$ to mean $\ulcorner$ if S is synthetic and atomic, S is empirically true. $\urcorner$ Then $\ulcorner P(S) \urcorner$ follows from S , $\ulcorner P(\sim S) \urcorner$ from $\sim S$, and so classical principles would produce $\ulcorner P(S) \vee P(\sim S) \urcorner$ which says in effect that every synthetic atomic sentence is determinate in empirical truth-value, a statement which is no more acceptable to us than the original principle of bivalence.

How should we respond to this? The fact is that the move from a sentence S to the assertion that S is empirically true fails for a more fundamental reason than that previously indicated, and this reason also applies in the case of the modified argument just stated. In the inference from S to $\ulcorner P(S) \urcorner$ a confusion is made between saying that certain evidence supports S , and therefore also $\ulcorner P(S) \urcorner$, and saying that S itself supports $\ulcorner P(S) \urcorner$. Acceptance of the first of these statements does not oblige us to accept the second. The statement that if S is synthetic and atomic, then the totality of evidence ultimately available (to anyone) favours S does not follow from S . It may follow (though probabilistically, not logically) from the assertion that some well-informed person has asserted S , but not from S

itself. Evidence, in our system, consists in the occurrence of relevant experiences and, in general what a given synthetic statement will justify is merely the belief that if there occur experiences of certain kinds, then there will probably also occur experiences of certain other kinds. There is no need to adopt principles that permit any more than this. Of course where S is a categorical experiential assertion $\ulcorner P(s) \urcorner$ does follow from S -- for in that case S entails the existence of the 'evidence' that is the very experience(s) that S categorically describe(s). But this is no problem, for determinateness of empirical truth-value does hold for categorical experiential sentences. If I am not mistaken then, the new argument for the thesis that our principles entail abandoning classical logic fails like the original one.

The ideas of the preceding paragraphs can be used as the basis for a brief comparison with the intuitionist attitude to classical logic. In the area of mathematics, the intuitionist holds, in effect, that from any proposition P one can derive the proposition that there is a proof of P and that from the negation of any P one can derive the proposition that there is a refutation of P. From this it is obvious that the intuitionist has to reject the Law of Excluded Middle, for it would commit him to the thesis that there was either a proof or a refutation of any

mathematical proposition. An extension of the intuitionist position from mathematics to factual knowledge might issue in the suggestion already examined that $\ulcorner P(S) \urcorner$ can be derived from S and $\ulcorner P(\sim S) \urcorner$ from $\sim S$. But we have seen that we can avoid accepting that notion, and so it is not necessary to follow the intuitionists in their abandonment of classical logic. Incidentally, there will be a comparison of our approach with another aspect of the general philosophical orientation underlying intuitionism when we come to a brief discussion of some of Michael Dummett's ideas in the last section of the final chapter (pp. 96-100).

3. THE DEFINITION OF EMPIRICAL TRUTH

In this section I attempt to define empirical truth. I should, however, warn the reader that only an approximate definition is achieved, since certain problems are raised but not solved. The suggestion made could, however, form the basis for further work.

As already indicated, a necessary condition for a sentence S to be empirically true is that it either have a high acceptability relative to U (the totality of experiences) or be satisfied by U. This is already an improvement on certain past attempts to define truth in terms of acceptability. Thus it has sometimes been suggested that a sentence is true precisely when it has a high acceptability -- meaning by this a high acceptability relative to all the evidence available. But clearly this will not do, since a sentence can be very well supported by all available evidence, and yet still turn out to be false.[5] This is why we require that an empirically true sentence be highly acceptable relative to all the evidence ever available to anyone, or rather, in terms of the present theory, relative to the set of all experiences that have ever existed or will ever exist. So the next suggestion is that S is empirically true just in case S has a high acceptability relative to U or is satisfied by U. But there is a serious problem with this suggestion. It might be the case that a disjunction $\ulcorner p \vee q \urcorner$

had a high acceptability relative to U but neither p nor q themselves had high acceptabilities relative to U. In such a case, the present suggestion would require $\ulcorner p \vee q \urcorner$ to be true,[6] but neither p nor q themselves to be true. The suggestion has to be amended in some way to take account of this.

It might be thought that a parallel problem could arise with respect to existentials. We cannot, it might be argued, call $'(\exists x)Fx'$ for example, 'true' in a case where no instantiation of $'Fx'$ is true, and this is just what the present suggestion might in some circumstances require us to do. Existential sentences are, on this way of thinking, 'parasitic' in some sense upon singular sentences in the same way as disjunctions are 'parasitic' upon their disjuncts. But this, I think, would be a mistake. Quine has shown that names, as a special category of expression, are redundant.[7] So the category of 'singular sentences', members of which are supposed to serve as instantiations of open sentences, is not a fundamental one. Singular sentences are themselves basically existential and so to make existential sentences parasitic upon singular sentences would be to make existential sentences parasitic upon other existential sentences, which would clearly distort the original intention. So the appearance of a problem here is perhaps illusory.

The first problem, however, remains. And there are others. Suppose S is true on the suggested definition of truth, i.e. has a high acceptability relative to U. And suppose S gives inductive support to a further sentence S', which, however, because of strong evidence against it, (i.e. against S') does not have a high acceptability relative to U. (There is no incompatibility between the presence of this strong evidence against S' and the high acceptability of S. The evidence against S' may not affect that acceptability. Just because S inductively supports S', this does not mean that $\sim S'$ inductively supports $\sim S$). It might be that if the acceptability of S relative to U were even higher than it in fact is, that evidence, when 'absorbed' by S', might be sufficient to counteract the evidence <u>against</u> S' thus causing it to finish up with a high acceptability relative to U, and therefore, according to the suggested definition, 'true'. But in fact if S is <u>true</u> then its <u>actual</u> acceptability relative to U should be no obstacle to the truth of S'. In considering whether or not S' is true, we should consider, not its <u>actual</u> acceptability relative to U, but its acceptability on the (false) assumption that S is as well supported as it possibly could be by U -- on the assumption, in fact that S is satisfied by U. In other words, our classifying certain sentences as true automatically boosts the prospects

for truth of other sentences that depend on them for support. This is something which our present definition of truth does not allow for.

There is a third problem. It turns out that on the present definition it is possible for a number of truths to entail a falsehood. Let α be a large set of sentences all with high acceptabilities relative to U. Suppose S is entailed by α. If the other conditions of the Rule of Entailment are satisfied, the support which all the members of α have will be transmitted to S, but there may also be so much evidence against S, that S actually has to be regarded as false. The latter is not incompatible with saying that every member of α has a high acceptability relative to U. For the strong evidence against S will weaken the conjunction of all the members of α, but, provided α is large enough, it will not sufficiently weaken any individual member of α. That is to say, although the strong evidence for the negation of the conjunction combined with the evidence for each of the individual members of α apart from some particular member, say A, will weaken A itself, it will not weaken A significantly, provided α is large enough.[8] Intuitively, each member of α can still be strongly maintained, because S's failure to be true can be blamed on a failure on the part of the others (with respect to each member, it can be

considered unlikely that it is <u>that</u> member that causes the failure of S).

The latter two problems I leave as something to be dealt with in future research. They do not look like the sort of problems that it would be impossible to solve. It would presumably be simply a matter of imposing certain further technical restrictions on what sentences are to count as empirically true.

The following modified definition of empirical truth is intended to take care of the first problem (pp. 58-9): S is empirically true just in case S either has a high acceptability relative to U or is satisfied by U and in addition there are sentences $T_1, T_2 \ldots T_n$ satisfying conditions (i)-(iii) below:

(i) Each T_i is either an atomic sentence within S or the negation of one.

(ii) Each T_i either has a high acceptability relative to U without S or is satisfied by U.

(iii) $\ulcorner (T_1 \& T_2 \& \ldots T_n) \supset S \urcorner$ is logically valid.

We define $\ulcorner S$ is empirically false$\urcorner$ as $\ulcorner \sim S$ is empirically true.$\urcorner$

Clearly if there are any defined expressions in the language, then the phrase 'atomic sentence within S' in the definition should be understood, in its application to a sentence using any of these defined expressions, to mean an

atomic sentence with the fully-analyzed version of S .

The idiom 'S is empirically true' is used when one has a particular language in mind. But there is also the dyadic predicate 'S is empirically true in L' where L ranges over languages, expressing a notion of empirical truth relativized to languages. This could be defined too, by replacing the notions of 'inferential factor' 'acceptability' and 'satisfaction' by new versions of these concepts relativized to sets of linguistic rules (for example, one would talk about an inferential factor relative to a certain set E of experiences and relative to a certain set R of linguistic rules). The explication of these new concepts would presumably offer no new problems, given the explications of their unrelativized versions.

CHAPTER FOUR : OBJECTIONS AND DEFENCES

1. GENERAL DOUBTS

One major doubt that one might have about our semantic method is that there seems to be no clear criterion of success in the finding of appropriate semantic principles -- or even a clear criterion of 'appropriateness.' With phenomenalism the situation appears very different. Success is judged by the ability to translate non-sense-datum statements into sense-datum statements. If the system fails to provide a translation for some non-sense-datum statement then to that extent it is incomplete; and if it provides a translation which does not accord with what intuitively seems to be conveyed by that statement, then it is to that extent incorrect. What is the corresponding criterion of success in the case of our method? In fact the appearance of there being no criterion is an illusion. Basically the system must produce intuitively correct judgments concerning the acceptabilities of sentences relative to sets of experiences. Of course we have no intuitions attributing numerically exact acceptabilities to sentences. The numerical values are fictions of the system. But we can see intuitively whether or not, for example, a given sentence is more acceptable relative to some set of experiences than to some other. What we require, then, is that our principles attribute numerical acceptabilities to sentences in accordance with

these judgments. If there is any discrepancy in this respect, the system must be considered defective (though it may be its inductive, rather than its semantic, part which is to blame). If one had a complete and detailed system this is the criterion that one would apply to it. Of course we have here scarcely begun to produce one. Arguably, only the achievement of this task together with good grounds for supposing that it met this criterion would constitute a really compelling proof that our method was workable. In this thesis I can only claim to have shown that it has some initial plausibility.

Another possible source of anxiety is the fact that the theory makes reference to entities other than experiences. A sentence, for example, can hardly be regarded as an experience. But here we should bear in mind certain important observations made in the first chapter. A semantic theory which is intended to be non-realistic with respect to entities apparently referred to by certain expressions of the language which it describes need not avoid the use of those expressions. For their use only involves a commitment to the entities in question if that use is representational in nature, and there is no reason why the non-representational explication of the expressions as they occur in the 'object language' should not be regarded as applying also to their use in the theory itself.

All this may appear dangerously liberal, for it

seems to countenance vicious circularities. Here we should refer back to the possible theoretical structure described on p. 11 in which physical concepts are explicated by axioms linking them to observational concepts, the observational concepts are explained in terms of 'stimulus-meaning,' and stimulus-meaning itself is ultimately understood in terms of the very physical concepts that the theory is intended to explicate. I wish to argue that while this is indeed incoherent its incoherence is not of a kind which characterizes every case in which a theory uses expressions which are amongst those which it is intended to explicate. Circularity itself is not to blame. It is, for example, surely impossible to produce an explication of the truth-functional connectives that does not actually use at least some of those connectives, for such expressions are absolutely basic to human discourse. Yet the explications that can be produced are not for all that considered defective.[1] What is wrong with the hypothetical structure described above is that it is <u>purely formal</u>. Ultimately it succeeds only in relating the use of certain expressions to the use of certain others, and so does not constitute a genuine explication of those expressions. Arguably for a semantic theory covering more than merely logical expressions to be adequate there must be at least some link with pure non-linguistic experience. Clearly the theory advocated in our second

chapter satisfies this condition in a fully explicit manner.

As I have already remarked in my first chapter, the theory whose inadequacy I hope I have just explained appears to be something like what Quine has in mind in his writings. Whether that is strictly so or not, it does not seem to me that Quine is a realist, whatever he might actually claim. Of course he does accept the Principle of Bivalence. What enables him to do this is his acceptance of a multiplicity of possible theories for the same totality of evidence.[2] Consider a sentence S which U (the totality of all evidence ever available) renders indeterminate in truth-value -- a possibility as real for Quine as it is for us, since he is the very originator of the doctrine of the underdetermination of theory by evidence (incidentally, when we say that U makes S <u>indeterminate</u>, we do **not** mean to imply that there is a real third truth-value involved -- merely that U fails to justify either S or $\sim S$). If we allow any theories that include either S or its negation, then we must have <u>both</u> theories that include S <u>and</u> theories that include $\sim S$, or else there will be a quite unacceptable arbitrariness. Quine accepts the consequent of this conditional in order to accept the antecedent, and so manages to maintain bivalence. But there is obviously another possibility. One can deny that U legitimizes any theory which includes either S or $\sim S$. This is our course, and it entails

rejection of bivalence. It is preferable to Quine's bacause if U does not (via whatever linguistic and inductive rules we are using) give S a truth-value, then if it is to have one it will have to be given one by some extra 'ad hoc' means. But there does not seem to be any good reason to do this. Certainly abandonment of bivalence constitutes no serious problem if, as we have seen in the last chapter, it does not necessitate revision of classical logic.

It should be pointed out however that, confusingly enough, we too accept that U is compatible with a multiplicity of possible theories. For U itself does not determine what kind of descriptive language we use and, as should be fairly apparent by this stage, we consider a language itself to constitute a theoretical structure. But the type of pluralism referred to in the previous paragraph is slightly different. Perhaps it should be regarded not as a pluralism of theories but as a pluralism of descriptions of the world in terms of some theory. Our basic theoretical structure is considered as already given, and what we are asking is whether it, in conjunction with U, allows for more than one possible description of the world (of a given degree of specificity). Our answer is no, since we do not allow our world-description to pronounce on those points which, being indeterminate, could create the plurality.

Quine's commitment to bivalence makes him look like a realist. But, as the foregoing should have made clear, the way in which he saves bivalence -- via a pluralism of theories -- is thoroughly non-realistic in character. One can say that the physical world is objectively real, but what significance is left to those words if one allows a multiplicity of possible physical realities? Again, the advocacy of a pragmatic standard in choosing between rival theories runs counter to realism. But the basic point is that Quine's conception of physical language is, as we have seen, non-represen tational, and this is incompatible with realism with respect to physical things.

2. DO 'EXPERIENCES' EXIST?

One aspect of the present approach that would undoubtedly occasion considerable oppostion is its very granting of independent ontological status to entities called 'experiences'. This is particularly likely in view of the fact that according to my conception of the matter, experiences include, as a proper subset, the much-loathed category of 'sense-data'. Technically, one might object to saying that sense-data themselves were experiences, preferring to talk of the experiences of having sense-data of which the sense-data themselves were the 'objects', but I do not see that there is anything to be gained by this distinction. Either way the existence of sense-data is a central element in the theory, and this requires defence.

The term 'experience' itself may be defined by giving examples. A pain, a thought, a feeling of fear, an after-image and a taste are all experiences. All these things are characterized by a certain 'immediate presentness' which may be regarded as the mark of experience-hood. Sense-data may be defined as those experiences that are associated with seeing, hearing or otherwise sensing things (a definition which is not intended to imply that whenever someone has a sense-datum something is 'genuinely' seen or heard etc. -- that the sense-datum has to be 'veridical'). There are,

broadly speaking, two kinds of attack on sense-data. One alleges that the very notion of these entities is somehow incoherent or confused (Austin and Wittgenstein are two examples of philosophers who take this line). The other argues that while it may be granted that sense-data do exist or that one can form a coherent notion of them, they do not have the importance which has traditionally been accorded them (many philosophers have argued for this -- it seems to be a much more common thesis than the first one). The present section will deal with the first kind of argument, the subsequent one with the second.

A fundamental fact about sense-data is that no sense-datum is identical to any physical thing or to the surface of any physical thing. One might try to argue for this by pointing out that sense-data may be present when the physical things that one might be inclined to identify them with are absent. For example, a sense-datum of a table may be present when there is no table present. But this may be countered by responding that it need not be asserted that <u>every</u> sense-datum of a table is identical to a table (or to the surface of a table) but only those sense-data that <u>do</u> happen to be veridical. However, there is another equally simple observation which I think clinches the case against the proposed identification -- the table may be there when no one is seeing, hearing or feeling it i.e. when there exists no sense-datum of it. This,

I think, makes it impossible to say that the table or its surface could be a sense-datum.

The strategy of our argument here should be carefully observed. We have not defined 'sense-datum' in such a way as to contrast sense-data with physical things, and so we have not begged the question in favour of the sense-datum-physical object distinction. Rather we have defined a sense-datum as a certain type of experience -- a type of entity which is immediately present to us -- and then found reasons for refusing to identify sense-data thus defined with any physical thing. We have thus established, in what I think is a non-question-begging way, that there is a type of immediate presentness which physical things do not have in relation to us as perceivers.

But it might be objected that we have been proceeding too quickly. While it may easily be granted that if there are such entities as sense-data they are distinct from physical things, one might question the assumption that we have to talk about our perception in terms of our immediate awareness of entities of this kind, or in general that talk about experience must take the form of talk of the existence of entities called 'experiences'. In other words one can object to what Quine calls illegitimate 'hypostatization'. One way of trying to show that the hypostatization involved in the sense-datum theory need not be accepted is to try and develop a way of

of talking about perception whose 'logical grammar' does not seem to require it. An example of this is the 'Adverbial' theory of perception which attempts to construe perceptual qualities not as properties of perceptual entities but as 'ways' of sensing. Thus, for example, this theory would replace 'I see a red sense-datum' by 'I see redly' or something similar.

There are some perceptual statements which it is hard to analyze in an adverbial way. Consider, for example, 'The left-hand image is fuzzier than the right-hand one.'[3] But although this presents a challenge to the adverbial theorist, it is not perhaps one which he cannot meet. Here is a possible construal of the statement: 'I am seeing in such a way as to suggest that there are two objects such that I am seeing the left-hand one more fuzzily than the right-hand one'. It may well be that the adverbial theorist can deal with all such problematic cases with sufficient ingenuity. So he may well be right when he claims that in describing our perceptions we do not have to use a logical grammar which presupposes sense-data. But it is still possible to maintain that as a matter of plain fact there <u>are</u> such entities as sense-data. Indeed one can argue that sense-data are paradigms of the real. Common-sense might maintain that it is everyday physical things that are paradigms of the real, but this impression may be lessened if the notion of sense-data is introduced and sense-data contrasted with physical things. Then one may be prepared to

concede that the 'concreteness' which appears to be the reason for attributing an indisputable reality to physical things is actually a characteristic not of the physical things themselves, but of their sense-data. Again, one must not think that because sense-data are entities, they are objects somewhat similar to physical objects. They do not have weight or volume, for example. Nor, oddly enough, do visual sense-data have color, in the sense in which this is attributed to physical things. As we have made clear on p. 39 the blueness of physical things must be carefully distinguished from the blueness of sense-data, the former being defined in terms of the latter. Failure to attend to points of this kind leads to erroneous criticism of the sense-datum theory. For example, R.M. Chisholm argues that since whiteness is the tendency to present a white appearance under favourable conditions, the whiteness of an <u>appearance</u> must be its tendency to present a <u>second-order</u> white appearance under favourable conditions and so on <u>ad infinitum</u>.[4] But this is refuted by the observation that it is only the whiteness of <u>physical</u> things which is analyzed as the tendency to present a white appearance. The whiteness of an appearance must be understood differently -- perhaps as unanalyzable. A similar argument to Chisholm's occurs in Gilbert Ryle where it is argued that since glimpsing something involves having at least one sensation and having a sensation is, according to

the sense-datum theorist, itself to be regarded as observing or glimpsing a certain kind of entity, having a sensation must involve having at least one *further* sensation of a higher order, and so on *ad infinitum*.[5] The response to this is that a plausible version of the sense-datum theory will not regard having a sensation as equivalent to *observing* a sensation. If I have a sensation this consists in the existence of a certain kind of entity which has as one of its properties the property of being an experience of mine. As we shall see in the last section, there is no reason to suppose that there are souls or egos and so this property of being an experience of mine need not be regarded as literally consisting in a relation to a further entity, but can be seen as unitary. But if we speak non-realistically as if having a sensation *did* consist in being an ego that stands in a certain relation to a sense-datum, we cannot maintain that this relation is that of observation. Physical things are observed, not sensations. The relation between perceivers and the physical things they perceive cannot be the same as that between perceivers and the sense-data that are immediately present to them, just because the two relations are *definable* only be reference to the type of object they involve and these two types of object are fundamentally different, as we have seen. Ryle is indeed at pains to point out that observation is not the same as

having a sensation. But I do not understand why he takes the negation of this statement to be essential to the sense-datum theory, unless he is merely relying on loose statements of that theory.

Another line of criticism of the sense-datum theory starts from the fact that our experiences are already pregnant with interpretation, and therefore, in a sense, with theory, so that there is no possibility of isolating a bare 'factual' substratum of knowledge which is simply apprehended without any interpretation being involved. Now I agree that there is no such substratum. But the experiences in which I believe are not intended to constitute such a substratum. I think that this part of the traditional theory of sense-data should be abandoned. For example, it does not seem right to say that our sense-data lack spatial depth, although just what their possession of this quality consists in is hard to say.[6] On the other hand there do seem to be aspects of the interpretation of sense-data that are not 'internal' to the sense-data themselves. An example is the belief that two sense-data separated in time are sense-data of the same object. Nothing in a pair of sense-data can mark them as standing in this relation in the same powerful way that a visual sense-datum presents depth.[7] For being of the same object as some other sense-datum is never part of the 'essence' of a sense-datum, whereas (it seems to me) having a certain kind of

depth can be (to put the matter in terms of our technical apparatus a statement asserting the latter can be satisfied by a set of experiences, but a statement asserting the former can only be more or less acceptable relative to a set of experiences). A careful distinction has to be drawn here between saying that a sense-datum is interpreted and saying that some of its properties are the results of interpretation. The quality of depth within a sense-datum can be said to be (at least according to a certain psychological viewpoint) the result of an interpretation. However, when sense-data are 'assigned to the same object' this interpretation does not affect the qualities of the sense-data themselves. It is an external operation performed on the 'already constituted' sense-data. The rejection of the 'uninterpreted substratum' really means the acceptance of sense-data some of whose qualities are the results of (or 'represent') interpretations -- sense-data that to some extent come 'ready interpreted', so to speak. Since it appears that there are such sense-data, the 'uninterpreted substratum' must indeed be rejected. But there is another respect in which traditional theories in this area seem to have been misguided. This is in their apparent conception of the nature of the 'external' variety of interpretation. It is not usually, as they often seem to suggest, a matter of conscious inference. I do not as a rule consciously infer the existence of objects from the

presence of sense-data. Rather the sense-data cause me to believe, in an immediate fashion, that the objects are there.[8] So a phrase like 'external operation' (used earlier) should not be taken too seriously (but note that the fault is with the word 'operation', not the word 'external'). However, this should not cause us to doubt that what is present to me in the most immediate way is the sense-data, not the objects themselves.

3. HOW IMPORTANT ARE 'EXPERIENCES'?

This section will focus specifically on the question of to what extent the world of sense-data constitutes the solid 'rock' upon which human knowledge is founded, as empiricists have traditionally maintained. We have already seen one respect in which this thesis would have to be weakened. The existence of sense-data that are partly 'ready interpreted' seems to require acceptance of a theoretical element already present in sense-data. Another issue relevant to this question is that of whether true statements about sense-data are incorrigible, that is to say whether they can be known for certain to be true, beyond any possibility of future revision. In The Problem of Knowledge A.J. Ayer points out that statements about my past or future experiences cannot be certain for me since it is always possible that my memory of a past sense-datum or my prediction concerning a future one may be faulty, and so the most that can be claimed is that I can be certain of statements about my present sense-data.[9] Ayer in fact goes on to argue that even here there is no significant incorrigibility, and I agree with him on this. Arguments for the latter point are intricate, but fortunately we do not need to concern ourselves with them. For what we have already observed is enough to undermine the view

that sense-data constitute an incorrigible foundation for knowledge. For as Ayer points out, a certainty which disappears as soon as the sense-datum that it concerns slips into the past is not outstandingly useful.[10]

But the abandonment of the incorrigibility thesis is hardly a disaster. Although statements about sense-data are never absolutely certain (as with all other factual statements, presumably), it may nevertheless still be claimed that they have sufficient certainty to serve as the justification of all factual claims about the 'external world' (and if we are talking of statements concerning experiences, and not merely sense-data, we can perhaps claim that such statements serve as justifications of all factual statements whatsoever).

But even this position can be attacked. We have already encountered A.M. Quinton's observation that our experiences tend more often to be the cause of our factual beliefs than the justification of them. This is connected with another point that Quinton makes viz. that close attention to our experiences is a comparatively abnormal occurrence and requires a highly specialized state of mind.[11] A different view, but one which is equally antagonistic to the theory of the epistemological priority of sense-data, is that of Nelson Goodman. He argues that it makes no sense to speak of my seeing a cardinal bird, say, either 'as' a cardinal bird or 'as' a red patch.[12] It is true that I may

come to believe that there is a cardinal bird present because I am aware of a red patch, but equally I may come to believe that I was aware of a red patch because of my knowledge that I saw a cardinal bird. Goodman believes that it is worthwhile to attempt the construction of a phemomenalistic system, but not because of the epistemological priority of phenomena, which he denies. He believes that there is no one absolutely correct constructional system, but that we can learn different things from different constructional systems. Are there any reasons to doubt Goodman's relativism? And if so, should we adopt the position that sense-data constitute the primary basis of justification or Quinton's view that it is physical things themselves that most often fulfill this role?

It is obviously true that not every justification ends in an appeal to experiences. If asked to justify my belief that Tokyo is the capital of Japan I might assert that it says so in a number of books. This does not of itself refer to any sense-data, but it is probable that if pressed further with the justification, I would be led to make such a reference. Thus if I were asked how I knew that these books asserted Tokyo to be the capital of Japan, I might reply simply that I had read them, thus invoking certain kinds of visual sense-data. What I wish to suggest, then, is that there is a certain <u>theoretical</u> incompleteness in many epistemic justifications which in a practical sense are perfectly adequate. And one can plausibly

maintain that any justification of a factual statement which is taken as far as it can be taken will end with a reference to some set of experiences, and it is this which constitutes the epistemic priority of experiences.

At one point Quinton considers whether it would make sense to ask for a justification for the statement 'This is a horse' as uttered by someone standing in front of a horse looking at it.[13] He suggests that the question 'How do you know?' would scarcely make sense, but that if pressed the person concerned would probably reply 'Well, because it looks like a horse,' which, however, should not be taken as genuinely expressing a reason but rather as representing 'an infusion of tentativeness into the original claim expressing a lack of confidence inspired by the nagging question'. The sense-datum theorist's interpretation of the reply as a description of the person's experience is rejected by Quinton as utterly implausible. But if the comment is not intended as a reason what is the word 'because' doing there? And if the sense-datum theorist's interpretation consists in the presentation of what would be a very good justification of the statement, this surely constitutes a good reason for supposing that the interpretation is correct. And it does appear that to say that one has in one's visual field a sense-datum of the type commonly associated with the presence of a horse constitutes a good reason for supposing that there is indeed a horse present.

Quinton offers no solid reason for preferring his analysis of the reply as an attenuation of the original statement over this analysis.

In general, there seems no reason to doubt that any justification for some factual matter concerning the 'external world' will, if taken far enough, lead back to something that somebody has seen, heard or otherwise sensed, where 'sensing' is to be understood in terms of the occurrence of sense-data. However, such a justification need not lead back to the detailed properties of sense-data (to suppose otherwise would necessitate rejecting Quinton's admittedly true observation that careful attention to our sense-data is comparatively rare). More often than not it will either lead back to very obvious features of sense-data (e.g. redness) or to the fact that certain sense-data are of the kind (whatever kind that may be) that is associated with the presence of certain objects (it is possible to recognize this association without being fully aware of detailed properties of the sense-data).

In this section and the previous one we have defended the sense-datum/physical thing distinction and tried to stress the importance of sense-data. In the following section we go further still and try to justify withholding ontological recognition from physical things altogether.

4. IS OUR ONTOLOGY TOO NARROW?

That physical things exist in some sense is not denied here. Physical thing talk, interpreted as 'logically parasitic' upon experiential talk, is recognized as legitmate, and to that extent physical things can be said to exist. But the point is that there is no reason to attribute to them the kind of full ontological status which we believe experiences to have -- the kind of full ontological status that is associated with representational linguistic explanations. It is to be noted that strictly speaking our position on this question is agnostic -- there may, for all we know, be entities that can be matched representationally with physical thing expressions. It is simply that we have no good reason to suppose that there are.

In the following section we shall adopt the analogical argument for the existence of other people's experiences. There does not seem to be any corresponding analogical argument that one could use to support a belief in the existence of physical things. What arguments could one offer? It might be suggested that it is reasonable to suppose that our sense-data have causes and since one cannot find these causes amongst the world of experiences themselves, they must be located in a realm beyond that of experience. But if there is any strong support for the statement 'For any x, something is the cause

of $\exists x$' (I do not say that there is), it seems to me that it can quite happily be taken as support for a non-representational interpretation of the statement and therefore not as showing that any event or state must be caused by some entity with full independent ontological status. In other words, one would need a further independent argument for supposing that physical language must be understood representationally if one wanted to use the fact (if it is a fact) that sense-data must have causes to support the idea that they must have literally real causes.

A further possible argument is that it is simply not rational to suppose that physical objects do not exist when things are exactly what they would be like if physical objects did exist. But the fact is that it is very misleading to say that things are exactly what they would be like if physical objects existed. What one should say is that things are exactly what they would be like if physical language were treated representationally and the entities thus represented did actually exist. The original statement presupposes a bias towards a representational view which is quite unwarranted. This argument is in fact fairly representative of a group of arguments which exploit the anti-realist's talk of 'useful fictions.' They could be summarized in the statement 'Surely if a fiction is useful, this constitutes a very good reason for supposing that it is really fact.' But I have already guarded

against presenting the anti-realist's views in terms of 'useful fictions' in my first chapter. To suppose that in using physical language we are 'pretending' that physical things exist is to presuppose that such language is representational. But there is no reason why we should consider it to be representational. And it is for this reason that any argument against anti-realism that depends upon attributing to the anti-realist the belief that physical language describes a fictional world will ultimately fail.

A different sort of argument for realism invokes the notion of 'simplicity.' It is suggested that a conception of the world which attributes independent ontological status to physical things is simpler and more unified than one which attributes it only to experiences, and therefore the former is to be preferred to the latter. But it seems to me that the hypothesis of fully real physical entities does not add significantly to the simplicity which is already involved in the recognition that experiences conform to regularities describable by non-representational physical statements.

A consideration which certainly makes anti-realism look very plausible is that there do seem to be actual theoretical indeterminacies recognized by scientists. For example, most scientists would agree that whether or not light is made of particles or waves is indeterminate. It is perhaps conceivable that further evidence might resolve this apparent indeterminacy, but what reason is there to suppose that it will? One might

try to respond to this point by claiming that the real physical world is not the world of modern science (or better, any of the possible worlds proposed by modern science), but rather the common-sense world of ordinary three-dimensional physical objects. But why should one take the view that a scheme which is so limited in its capacity for enabling us to predict and control experience in comparison with certain other possible schemes should be the one that actually represents reality?

It might be urged that although we may have no good rational support for a belief in physical things, nevertheless it is psychologically almost impossible to give up that belief. Now in the course of our everyday lives it is presumably psychologically impossible to avoid thinking of physical-object-expressions as representing entities of some kind and regard them instead as formal 'uninterpreted' symbols. But anti-realism with respect to physical objects does not require us to do that. It simply makes the point that as far as their efficacy in enabling us to predict and control our experiences is concerned, they might theoretically be regarded in that way, and moreover that this efficacy is indeed all that really matters about them.[14] But some might doubt that this is all that really matters about them. Might it not be that physical objects are emotionally significant in themselves, independently of their connection with our experiences (in the same way that other people's experiences are emotionally important to us, a fact

which makes it difficult to accept an anti-realist account of them -- see the following section)? For example, an astronaut might strongly desire to go to the moon. A certain physical object -- the moon -- is an essential constituent in this desire, it might be said. It is not enough for the astronaut to have all the experiences of going to the moon; he wants to really go there. But if he is not satisfied with having all the experiences of going there, what more is it that he is actually demanding? Perhaps that the entire course of his experiences throughout his life be consistent with his having gone there and that the experiences of other relevant persons be consistent with this too. In effect, he wants the universe of experiences to be such as to render true the statement that he goes there. So there is a sense in which physical things figure essentially in our desires, but oddly it is a sense which actually confirms the primacy of experience -- we want our experiences to be those that would arise in a world in which such-and-such physical states of affairs obtained. I therefore think that it is possible to maintain that, when all is said and done, our universe is quite rich enough if it contains only colors, sounds, smells, feelings, thoughts and so on.

5. IS OUR ONTOLOGY TOO BROAD?

To ask whether a scheme as puritan as the present one has too broad an ontology might seem absurd. But when we say that our theory commits itself to the existence of experiences we include the experiences of other people besides ourselves. The question then arises of whether we can justify this commitment. We must ask also to what extent we can justify belief in the future experiences of ourselves and of other people and even in past experiences.

The 'Other-Minds' problem has frequently been argued to be no more than a pseudo-problem. One could express the Wittgensteinian justification of this attitude by saying that other-minds-talk is logically parasitic upon behavioural talk, that in some sense it adds nothing to the behavioural talk, so that provided we have no doubts about the existence of the behaviour we should have no doubts about the existence of the experiences either. According to this view, the relationship between other-minds-talk and behavioural talk is essentially the same as the relationship that we are maintaining exists between physical-thing talk and experiential talk (in both cases the parasitic link is not of a kind to justify reduction). Completed in such a way as to fit our own standpoint, the argument for other minds becomes this: other-minds-talk is parasitic upon behavioural talk and

behavioural talk, being a species of physical-thing-talk, is itself parasitic upon talk of one's own experiences. Hence, having no doubt about the existence of my own experiences, I should have no doubt about the existence of other people's either.

It should be clear where this argument goes wrong. There may indeed by a sense in which showing that discourse about As is logically parasitic upon discourse about Bs, when there is no doubt about the reality of Bs, demonstrates the reality of As. For one thing it provides a response to philosophical scepticism about As by legitimizing A-talk by reference to B-talk. It answers the question 'How can we be sure, on the basis of our knowledge of Bs, that the statements we make about As are by and large true?'[15] But it is a major contention of the present work that there is also another important sense in which to reveal a relationship of this kind is to demonstrate the unreality of As. In a deeper, philosophical sense As have been shown not to exist. They have no independent ontological status. But it seems to me that is is precisely this sense of 'reality' which is at issue here. I consider it almost intuitively self-evident that for other people's experiences to be significant to us -- for other people's pains to exist in a way which makes sympathy an appropriate reaction, for example -- they must have full independent ontological status.

We must therefore ask whether one can have reasons for believing in the reality in this sense of other people's experiences. I think that the traditional inductive or

analogical argument may be used here. I can reason that since I am usually in pain when I behave in certain ways, then probably so are other people when they behave in those ways. The argument is not as strong as one would perhaps like it to be, for it postulates the existence of a whole new category of entities (experiences which are not mine) different in a fundamental respect from those already acquainted with (experiences which are mine). I am supposing that when someone else exhibits pain-behaviour there usually exists something analogous to what usually exists (viz. my pain) when I exhibit pain-behaviour. Such an analogical argument is much weaker than, for example, a straightforward inference from the fact that A's have always been accompanied by B's in the past to the conclusion that this A will be accompanied by a B. Nevertheless, it perhaps has some force.

But some might argue that we should not even get as far as considering possible inductive arguments for the independent reality in our sense of other people's experiences because the hypothesis is a meaningless one (such objectors would not assert that they alone have experiences but would probably adopt the Wittgensteinian theory of logical parasitism, repudiating our belief that this does not give them other people's experiences in any significant sense). For independent ontological status for other people's experiences

seems to involve a rejection of all logical links between experiences and bodies (since, as we have seen, it requires rejection of complete logical parasitism and, although a partial logical connection might be a theoretical possibility, it is hard to see what form the latter could take). And Strawson for one has argued that an essential component in the numerical identity of an experience is whose experience it is, in the sense of which body it is associated with.[16] Thus it would appear that the conception of other people's experiences that we have argued to be required is one that would make it impossible even to numerically identify an experience. But suppose that while agreeing that experiences must be identified partly by whose experiences they are, we refuse to construe this in terms of association with particular physical bodies. Then of course it is incumbent upon us to explain in terms of what we do understand the notion of an experience's being had by some particular person. It would be possible here to invoke Cartesian souls, but it does not seem necessary to do so. For since singular terms referring to Cartesian souls and variables ranging over Cartesian souls could presumably be limited to occurrence within the context 'is an experience had by _____,' the latter could be regarded as a unitary phrase. Hume seems to be right in thinking that my self, over and above the experiences it has, is a chimera. In a solipsistic world this means that no reference to persons is needed, since

the only person that could be referred to is myself and Hume's point shows that this is not necessary. In our broader world, it implies that although there must be apparent reference to persons in the idiom 'is experienced by ____,' the latter should not be taken as literally asserting a relation between an experience and a person but merely as expressing an unanalyzable property of experiences. But, it will be urged, the essence of our original problem still remains. It may now be expressed by asking what kind of property is involved here. My response to this will seem unsatisfactory, but I think it is ultimately defensible. I think it can be maintained that any workable world-view must include a reference to properties or relations the nature or 'content' of which we have no understanding of, and that the properties that we are dealing with here belong to this category. I shall refer to such properties and relations as 'inscrutable.' The countenancing of them may appear highly objectionable, but if I can indeed argue that any plausible system must recognize at least some inscrutable properties or relations, then the objectionable aspects of doing so will either have to be tolerated or shown to be illusory. It seems to me that certain spatial and temporal predicates of a type which must figure in any acceptable world-view can plausibly be argued to be inscrutable in my sense. Consider the assertion that a certain experience occurs after a certain

other experience. I do not think that I can say that I have any understanding of the content of such an assertion. I can picture the 'afterness' involved on the analogy of a spatial dimension, but such picturing is only picturing and the analogy is only an analogy. It does not represent the true nature of the relation. It may be argued that when two experiences of the same person occur very close together in time so that they are as it were enveloped in a single piece of consciousness, the person concerned has an immediate apprehension of the nature of the temporal relation of succession. This may or may not be so, but it clearly does nothing to throw light on the case where the two experiences are far apart in time. It should be noted that what we have here is not mere 'ineffability' such as might be held to characterize the meaning of a predicate like, say, 'red'. It is not merely that we cannot define or describe what is involved; we do not even have any intuitive understanding of it.

There may be some systems in which the notion of one experience's following another is not taken as basic. In such systems an assertion of this kind might be subjected to a truly reductive process which issues in a translation in terms of the primitives of the system. Or alternatively, the relation to the primitives might be of the weaker kind describable as 'logical parasitism'. But an obscurity of the kind we are

dealing with here seems too deep-seated to be removed by such methods. If they were used, the obscurity would surely reappear in at least some of the primitives of the system. So I think we can conclude that attribution of experiences to persons do not add a defect to any system which it would not already have in some form. Of course, it may be argued that we should have as few inscrutable properties as possible, but this is a much weaker argument, particularly in view of the fact that the inscrutable properties in question do, as we have seen, have some inductive grounding, and that without them we cannot really avoid solipsism.

Before we leave this part of the argument there is something that should be said about inscrutable properties which will hopefully be somewhat reassuring. This is that it is not the case that we cannot say anything about them. In the case of 'temporally succeeds', for example, we can at least say that this relation is transitive, asymmetric and irreflexive. In the case of 'is an experience of $\underline{a}$' we can say, for one thing, that other properties of experiences having this property are closely correlated with the properties of a particular physical body (this is a matter of definition -- the property of being an experience of $\underline{a}$ is introduced precisely as a property of entities whose postulated nature is to bear such a relation to a particular physical body). But it is clear that such facts about inscrutable properties

do not exhaust their content -- and this is what makes them inscrutable.

Another line of attack on our position might be derived from the ideas of Michael Dummett. Although he has never positively embraced a (controversial) anti-realistic position, Dummett has spent much of his career showing how it is possible to argue for anti-realism in many spheres. It might therefore be thought that his view-point would have a lot in common with that of the present work. But in fact if Dummett's arguments are to be taken seriously, they would seem to imply a far narrower ontology than even I am prepared to tolerate. For Dummett's idea is that no statement can be regarded as being objectively true or false that we could not in principle recognize as being true or false. Now it is clear that on my conception other people's experiences are not such that I can recognize them as obtaining or not obtaining. Hence Dummett's argument would lead one to take an anti-realist view of other people's experiences understood in this manner.

So what is Dummett's argument? His first point is that acceptance of realism with respect to some particular sphere entails acceptance of a truth-conditions theory of menaing for sentences apparently referring to entities within that sphere. This means that if a speaker conceives some category of entities realistically and if 'F' is a one-place

predicate which he and other speakers apply to some members of this category, a representation of what this speaker knows in knowing how to use 'F' will consist in sentences of the type '⌜a⌝ is F' is true if and only if a is F'. Dummett then points out that such knowledge must be manifested in some way and he argues that it can only be manifested by an ability to recognize whether or not the truth-condition obtains -- that is, to go through a decision-procedure resulting in a 'yes' or 'no' conclusion as to its obtaining -- and classify the sentence as 'true' or 'false' accordingly. Consequently a truth-conditions theory seems quite inappropriate for sentences -- like those concerning infinite totalities, for example -- which cannot be recognized as being true or false. Hence one cannot take a realist view of what such sentences concern.[17]

We cannot possibly accept the conclusion of this argument without revising our theory for it would require us to take an anti-realist view of many things about which we are realists. We have already mentioned that it would preclude us taking a realist view of other people's experiences. But in fact the divergence is even greater than this, for as Colin McGinn points out,[18] Dummett's view would even require us to give up a realist conception of our own past experiences.

McGinn has provided a plausible answer to Dummett's

argument.[19] He agrees that knowledge of realist truth-conditions must be manifestable in some way, but argues that it need not be manifestable by a capacity to recognize these conditions as obtaining. He argues that provided one can have a conception of states of affairs that transcend experience, there is no reason why knowledge of T-sentences should not be manifestable simply by a capacity to correctly interpret sentences as referring to such states of affairs. And the belief that one cannot acquire such a conception is, according to McGinn, just a dogma.

I agree with this response of McGinn's. My anti-realism with respect to physical things is not founded on Dummett's radical approach, but on the considerations raised in the previous section. Another consequence of this divergence between Dummett's approach and the present one consists in a difference of attitude towards infinity. Dummett will not accept a realist account of statements that quantify over an infinite domain. This is reminiscent of mathematical intuitionism, and Dummett himself explicitly maintains that the argument used to support it is a generalization of the intuitionists' central argument to areas beyond mathematics.[20] This is the argument that truth-conditions must be recognizable, which is given extra support, as we have seen, by Dummett's reference to the manifestation of linguistic knowledge. And it is clear that the argument would entail rejecting a realist account of sentences quantifying over infinite totalities,

because their truth-conditions, if such existed, would not be capable of being recognized as obtaining or not obtaining.

I shall not embark on a detailed defence of my position on this point, but it does seem to me that, for example, the statement that a reiteration of experiences of a certain type lasts forever makes sense in the straightforward kind of way that a realist interpretation requires. So in our system we should want this statement to be empirically true, if at all, not in virtue of its acceptability relative to U, but in virtue of its actually being satisfied by U. There is nothing to stop U, as the totality of all experiences, from containing an infinite number of members. This means that even if there were a super-being able to survey the entire contents of U, he might not be able to calculate the empirical truth-value of every sentence that had an empirical truth-value. Sentences dealing with infinite sets of experiences would constitute a problem for him. He could be aware of each member of such a set at some time i.e. of each experience relevant to determining the empirical truth-value of the sentence in question, but for this being to actually make the determination he would have to be aware of all of them at the same time, which is impossible. For Dummett, however, the super-being would be omniscient, since according to his (tentative) view, there are no truths that could not in principle be recognized as true. In contrast, we accept

the possibility of a truth which could not be known to be true, although it must always be as it were a function of truths which could.[21] A statement about an infinite set of experiences may be regarded as an infinite conjunction of statements all of which could in principle be recognized as true or false by someone. This observation could in fact be used as an argument for our position as against Dummett's. For how could a conjunction (albeit an infinite one) of statements all of which had an objective truth-value itself fail to have an objective truth-value?

We have given indications of how one can partially justify a belief in other people's experiences. Can justification also be given for belief in other experiences not immediately present to one, such as one's own future and past experiences? There is a fairly obvious inductive argument for the existence of one's future experiences based on the fact that one's experiences are associated with the state of a particular body and that one has good reason to suppose that that body will continue to exist in the future at least for some time. But this argument presupposes the reality of the past experiences that support these beliefs. Thus rational belief in the future seems to depend upon belief in the past. I shall not pursue the question of whether one can justify the latter. It seems likely that our minimum 'bearable' ontology includes more than can be shown to be rationally defensible in the strict sense.

NOTES

CHAPTER ONE

[1]In Quine, From A Logical Point of View, pp. 1-19.

[2]Ibid. p. 13.

[3]Ibid. p. 13.

[4]Of course, we have not mentioned Quine's arguments for his criterion. Anyone who indeed thought that 'Some species are cross-fertile' really does commit one to the existence of cross-fertility would presumably want to know the justification for the criterion. But in fact it is via the consideration of such questions as whether the use of a predicate commits one to a belief in properties that Quine is led to adopt his criterion. His reason for answering 'no' to this particular question is, in effect, that he cannot see what good reason one could have for giving an affirmative answer. One possible reason - that the use of an expression always involves naming something, and hence that a predicate must name a property - is disposed of in the course of his paper.

[5]The idea of such a program is presented in Davidson, "Truth and Meaning",

[6]The reader will almost certainly be aware that we are talking here about a project which many of the logical positivists advocated. See Hempel, "Problems and Changes in the Empiricist Criterion of Meanings", p. 179.

[7]See Quine, Word and Object, pp. 32-33. I derive the idea of this type of approach to meaning (not just the concept of stimulus-meaning) from Quine. It seems to represent the kind of theory that Quine himself would like, as indicated by "Two Dogmas of Empiricism" (pp. 20-46 of From a Logical Point of View) and the early chapters of Word and Object.

[8]Though one should add here the comment that the sense of the word 'real' associated with representational theories is the one that seems relevant to the philosophical problem of ontology. It is the only kind of reality worth getting worried about.

[9]We have given some reasons for doubting that

'formalist' theories of meaning do not involve us in ontological commitments. But this does not prove that representational ones do. One could, however, perhaps try to address particular doubts. Consider the following: a Tarski-Davidson theory for, say, English is a set of stipulations intended to generate the totality of sentences of the type '"Snow is White" means in English that Snow is white.' But the latter are trivial facts that no one can dispute. How then can they involve us in anything so serious as an ontological commitment? The answer is that it is an essential presupposition of the appropriateness of giving a Tarski-Davidson theory for a language that when one is told that "Snow is white" means in English that snow is white, or alternatively (to remove the similarity between object language and metalanguage, which is a red-herring here) that '"Snow is White" signifie en anglais que la neige est blanche' one is being told something significant, that this is the way to explain what the meaning of a sentence is. But it is hard to interpret that claim without supposing that we are being presented with a 'picture' theory of meaning on the lines of say, Wittgenstein's 'Tractatus' - a type of theory that does seem to involve us in genuine ontological commitments.

The reader may perhaps have hoped for a more exact characterization of the difference between representational and non-representational theories in this chapter. Unfortunately, this is very difficult to achieve. Perhaps the easiest way of showing what is intended is to give examples of representational theories. These would include the theory of language presented in Wittgenstein's 'Tractatus', that of R. Carnap in his Introduction to Semantics, that of N.L. Wilson in his The Concept of Language (a development of Carnap's views) and of course, Davidson's theory. I am not suggesting that there are not crucial differences between these various theories, but at the same time there do seem to be underlying similarities that justify grouping them together and considering them all to have the ontological implications I have attributed to 'representational theories'. One thing that all of them seem to involve is some kind of correspondence between linguistic units on the one hand and types of non-linguistic entities on the other. This is no more than an impressionistic remark. However, it may help us to clear up a certain misunderstanding. Strictly, saying that an explication of 'φ' is representational does not settle precisely what ontological commitment is involved in the use of 'φ'. This depends on what 'φ' is actually construed as representing - what precisely 'φ' is thought of

as 'corresponding to' in the extra-linguistic realm. If it is taken as representing ϕ-hood (as it might be in Wilson's theory, for example) then the commitment is to ϕ-hood. If, on the other hand it is taken as representing ϕ_s - things that are ϕ (as it might be in Davidson's theory where one might have a clause of the form 'an object satisfies 'φ' if and only if it is ϕ ' in which case the correspondence might very well be one-to-many), the commitment is precisely to ϕ_s. Thus when in the text it is said that an ontological commitment to ϕ_s is involved in a theory that entails '$(\exists x)\phi x$' when 'φ' is to be explicated representationally, this must be understood strictly as referring to a <u>particular kind</u> of representational explanation.

CHAPTER TWO

[1]I could, following writers like Michael Dummett, use the term 'assertability' instead of 'acceptability', except that 'assertability of *S* ' should mean 'extent to which it would be correct to assert *S* ' and this, strictly speaking, depends on other factors in addition to the likelihood of *S*'s being true (for example, whether *S* is relevant to the matter in hand).

A more important point is that our notion of acceptability relative to a set of experiences is only applicable to what Quine calls 'eternal sentences' i.e. context-independent sentences like 'The object H.J. Simmons points to at 5.00 p.m. on October 12th, A.D. 1982 is red at that time,' not context-dependent sentences like 'This is red.' The acceptability of sentences of the latter kind must be taken as relative not just to experience-sets (which render them more or less 'likely' given their precise reference) but also to contexts (which give them their precise reference). Such sentences would therefore presumably have to be dealt with by means of some kind of theoretical structure grafted onto the main structure for dealing with eternal sentences. It should be noted that for the purpose of 'explicating concepts' an exclusive concern with eternal sentences is not unreasonable.

[2]There are a number of terms used in this chapter, including 'acceptability', which, strictly, should always have 'relative to a set of experiences', or something similar, attached to them. However, I often use them elliptically without any such phrase, intending the reader to fill in for himself the appropriate additional words.

[3]By 'characteristics' we really mean characteristics not logically connected with relations which their possessors may bear to non-members of the set. Without this qualification almost any information about the world of experiences could be considered as bound up within any particular set of experiences. For example, if a certain experience *a* within a set *E* is blue and if we count as one of the characteristics of *a* the property of being the same in color as some other experience *b*, not within *E*, the acceptability of some *p* relative to *E* might be affected by the fact that some non-member of *E* viz. *b* is blue, which is contrary to our intention.

[4]The word 'render' should be taken seriously here. An examination of any set of experiences will leave one certain of a statement of the form $p \vee \sim p$ for example, but it is not the examination that renders it certain. The relationship involved is stronger than either material or strict implication. So logically valid statements are not satisfied by sets of experiences on this approach (nor can they have acceptabilities relative to sets of experiences). What makes such a statement logically valid is not satisfaction by every or any set of experiences but rather its following from some stipulated set of logical axioms.

[5]It might be wondered why we do not cover the case where every member of α is a theorem. In such a case S too would be a theorem, but it should be remembered that we do not want theorems satisfied by sets of experiences (see note 4).

[6]This definition already presupposes that α and S satisfy the conditions already stated viz. that α entails S, S is not contained in α, and that situation (B) (p. 29) obtains in relation to α and S.

[7]In our example on pp. 30-31, β is $\{p \supset q, q \supset r\}$ and p is $p \supset r$.

[8]In our example on p. 31, γ is $\{p \supset q, q \supset r\}$ and λ is $p \supset r$.

[9]In our example on pp. 31-32, R_1 is $p \& q$ and R_2 is p.

[10]In our example on p. 34, ω is $\{p, p \supset q, q \supset r\}$, β is $\{p \supset q, q \supset r\}$ and P is $p \supset r$.

[11]In our example on pp. 34-35, γ is $\{p \supset q, q \supset r\}$ and Q is $p \supset r$.

[12]This is emphasized, for example, by A.J. Ayer in The Problem of Knowledge, Chapter 3, Section (vii). The whole orientation of this chapter owes much to the discussion found here. Ayer sees clearly that it is possible to maintain that the meaning of physical language is to be explained in terms of sense-data without being a phenomenalist in the strict sense.

[13]Urmson, Philosophical Analysis: Its Development Between the Two World Wars, p. 156.

[14]The reader will very likely be unconvinced by what I have said about causal implications. Let me try to adopt a more forceful approach. Let us use the term 'verificationism' with respect to a particular class of expressions to refer to the thesis that one can give an adequate account of the use of those expressions simply be describing (a) the kinds of observation that justify their application and (b) the kinds of observation that can be inferred from their correct application (the apparent reference to 'possible observations' here is superficial. What we would really invoke are conditionals of a type similar, if not identical, to the very category of causal implications whose analysis we are discussing). Let us use the term 'actualism' to refer to the thesis that causal implications tell us only about the actual course of events in the world (this includes future states of the world) -- that they do not tell us about possibilities or possible worlds. Clearly if verificationism is true with respect to causal implications, then actualism is true. This follows straightforwardly from the fact that our observations can only be of the actual, not of the possible. But the crucial question is obviously whether verificationism is true with respect to causal implications. My suggestion is that the only good reasons one can ever have for resisting verificationism with respect to any class of expressions will be essentially metaphysical in nature. That is to say, they will be based on the fact that one has a commitment to certain para-observational entities -- perhaps because one thinks that there are good reasons for believing that such entities exist, or merely that one finds it impossible not to believe in such entities. The kinds of reason that I do not think that one can have for rejecting verificationism are those connected with the ideal of pragmatic adequacy -- arguments to the effect that the verificationist account does not do justice to the actual use of the expressions in question. For how could a verificationist account, being of the very nature that it is, fail to do justice to that (the arguments of the early part of Quine's Word and Object are perhaps relevant here. See McGinn, "Truth and Use", pp. 32-3)? If I am right about this, then the non-actualists will have to come up with some good metaphysical reasons for, say, believing in possible worlds, before we can take their approach seriously. In the absence of such good reasons, I think we should opt for actualism, despite its technical difficulties.

[15]Quine, From a Logical Point of View, pp. 65-79.

[16]Wilson, "Space, Time and Individuals," pp. 589-98.

[17]Putnam, Mind, Language and Reality, p. 227.

[18]The phrase 'semantic theory' seems a convenient way of referring to a theory for determining the acceptabilities of sentences relative to sets of experiences. However, it should be remembered that such a theory contains as a crucial element rules which have more to do with confirmation theory than with semantics.

[19]This obviously presupposes that we do have some way other than mere recognition of linking sense-data with particular objects. One way of establishing a given link of this kind might be by observing that it is the simplest hypothesis available under the circumstances.

CHAPTER THREE

[1]This follows the same broad pattern as Wilson's definition of atomic truth in The Concept of Language, p. 91. Of course, apparent reference to propositions is ultimately to be explained non-representationally.

[2]In fact we make it the case that no theorems are empirically true. It is only strictly necessary to deprive a theorem of empirical truth in the kind of case discussed in the text, but it is more convenient to withhold it from all theorems. Hence the statement made above that a necessary condition of empirical truth was having a high acceptability relative to U or being satisfied by U was correct (we did not need to add 'or is a theorem').

[3]This is reminiscent of the 'performative theory' of truth. See, for example, Strawson, "Truth".

[4]Note that the converse does hold. S does follow from $\ulcorner S$ is empirically true$\urcorner$ (at least if the object language is part of the metalanguage).

[5]The suggestion mentioned would, as we have seen, be more appropriate for our notion of redundancy truth. However, one of the reasons why we need empirical truth in addition to redundancy truth is in fact to explain why 'The evidence now favours S, but S is false' is not self-contradictory (even though it may be self-defeating). The explanation is that the falsity referred to is empirical falsity, not redundancy falsity.

[6]Throughout this section 'true' is to be taken to mean 'empirically true'.

[7]Quine, From a Logical Point of View, pp. 7-8.

[8]The mathematical probability of the occurrence of all of a large number of independent events each having the individual probability 0.75, say, will be much smaller than 0.75. Now it is true that our acceptabilities are not mathematical probabilities, but one would expect them to conform to a similar pattern.

CHAPTER FOUR

[1]These considerations suggest that explications in the present sense differ from conventional definitions in that they do not have to be capable of being understood by someone who does not understand the expressions they explicate. I owe these comments to Mr. B.F. McGuinness.

[2]This point comes from Colin McGinn's "Truth and Use", pp. 32-33, except that McGinn, connecting acceptance of bivalence with realism, states that it is realism itself that Quine's pluralism of possible theories allows him to maintain.

[3]The example is from N.J. Griffin, "Sense-data" (unpublished). Griffin actually uses it against a different kind of adverbial theory according to which perceptual qualities should be analyzed in terms of the way things look (i.e. common-sense things, not images). I think that the example is indeed more effective against that than against the type of theory discussed in the text.

[4]Chisholm, Theory of Knowledge, p. 95

[5]Ryle, The Concept of Mind, p. 203

[6]One thing that it clearly does not consist in is the literal possession of a third dimension. For a real third dimension in sensations would presumably be incapable of simulation in terms of the 'first two' as in a picture. Depth is not a real dimension just because it is marked out as separate from the 'first two'.

[7]Of course, this quality of visual depth should be carefully distinguished from the real physical depth which the object corresponding to the sense-datum may or may not have and of which the visual depth is a sign. This 'real depth' is more like the case of two sense-data belonging to the same object.

[8]Quinton, "The Problem of Perception", p. 70.

[9]Ayer, The Problem of Knowledge, p. 58.

[10]Idem.

[11]Quinton Op. Cit. p. 68.

[12]Goodman, The Structure of Appearance, p. 101.

[13]Quinton Op. Cit. pp. 80-81.

[14]The question arises, not just of whether it would be possible to give up thinking of physical objects as literally real, but also of whether it would be possible to give up using physical object language altogether. It seems to me that this would be possible for beings pyschologically and cognitively constituted very differently from us (who could, for example, give very accurate descriptions of their experiences). I also think that it would be possible in principle to develop rules for predicting future experiences on the basis of past ones that referred only to experiences, but presumably there would be no straightforward isomorphism between such rules and the structure of physical language, since otherwise straightforward reduction of the physical to the phenomenal would be possible, which, as we have seen, it is not.

[15]Ayer Op. Cit. pp. 144-148.

[16]Strawson, Individuals, chapter Three.

[17]One source for Dummett's argument if "What is a Theory of Meaning? (II)" in G. Evans and J. McDowell (ed.), Truth and Meaning.

[18]McGinn Op. Cit. p. 36.

[19]Ibid. p. 30

[20]Dummett, Truth and Other Enigmas, p. 17.

[21]Ibid. p. xxvii.

BIBLIOGRAPHY

Ayer, A.J. The Problem of Knowledge, MacMillan, 1956.

Carnap, Rudolf Introduction to Semantics, Studies in Semantics, Vol. One, Harvard University Press, 1942.

Chisholm, R.M. Theory of Knowledge, Prentice-Hall Inc., 1966.

Davidson, Donald. "Truth and Meaning", Synthese, 7 (1967), pp. 304-23.

Dummett, Michael. Truth and Other Enigmas, Harvard University Press, 1978.

----------. "What is a Theory of Meaning?(2)" in G. Evans and J. McDowell (ed.), Truth and Meaning, Clarendon Press, 1976.

Goodman, Nelson. The Structure of Appearance, third edition, Boston Studies in The Philosophy of Science, Vol. 53, D. Reidel, 1977.

Griffin, N.J. "Sense-data" (unpublished).

Hempel, Carl. "Problems and Changes in the Empiricist Criterion of Meaning", in L. Linsky (ed.), Semantics and the Philosophy of Language, University of Illinois Press, 1952.

McGinn, Colin. "Truth and Use" in Mark Platts (ed.), Reference, Truth and Reality, Routledge and Kegan Paul, 1980.

Putnam, Hilary. Mind, Language and Reality, Philosophical Papers Vol. Two, Cambridge University Press, 1975.

Quine, W.V.O. From a Logical Point of View, second edition, Harvard University Press, 1980.

----------. Word and Object, M.I.T. Press, 1960.

Quinton, Anthony. "The Problem of Perception" in G.J. Warnock (ed.), The Philosophy of Perception, Oxford University Press, 1967.

Ryle, Gilbert. The Concept of Mind, Penguin, 1963.

Strawson, P.F. Individuals, Methuen, 1959.

----------. "Truth", Analysis, vol. 9 (1948-9), pp.83-97.

Urmson, J.O. Philosophical Analysis: its development between the two world wars, Oxford University Press, 1967.

Wilson, N.L. The Concept of Language, University of Toronto Press, 1959.