1988

The Infirmities Of Externalism: Putnam And Dummett On The Realist Programme In Metaphysics And The Philosophy Of Language

David Alan Davies

Follow this and additional works at: https://ir.lib.uwo.ca/digitizedtheses

NOTICE

AVIS

The quality of this microform is heavily dependent upon the
quality of the original thesis submitted for microfilming
Every effort has been made to ensure the highest quality of
reproduction possible.

If pages are missing, contact the university which granted
the degree.

Some pages may have indistinct print especially if the
original pages were typed with a poor typewriter ribbon or
if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, pub-
lished tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed
by the Canadian Copyright Act, R.S C 1970, c C-30

La qualité de cette microforme depend grandement de la
qualité de la thèse soumise au microfilmage Nous avons
tout fait pour assurer une qualité supérieure de reproduc
tion

S'il manque des pages, veuillez communiquer avec
l'université qui a conféré le grade

La qualité d'impression de certaines pages peut laisser à
désirer, surtout si les pages originales ont été dactylogra
phiées à l'aide d'un ruban usé ou si l'université nous a fait
parvenir une photocopie de qualité inférieure

Les documents qui font déjà l'objet d'un droit d'auteur
(articles de revue, tests publiés, etc) ne sont pas
microfilmés

La reproduction, même partielle, de cette microforme est
soumise à la Loi canadienne sur le droit d'auteur, SRC
1970, c C-30

THE INFIRMITIES OF EXTERNALISM: PUTNAM AND DUMMETT

ON THE REALIST PROGRAMME IN METAPHYSICS AND THE

PHILOSOPHY OF LANGUAGE

by

David A. Davies

Department of Philosophy

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario.
August 1987

Part One, which is primarily exegetical in intent, traces the development of Putnam's thinking on our themes from the pre-critical "sophisticated realism" that culminates in the Locke Lectures, to the internalism developed in Reason, Truth, and History and later works. Chapter One examines Putnam's treatment of "realism" and related notions in the period prior to the 1976 Jerusalem conference on the philosophy of language, where - too close to Damascus to permit one to entirely avoid recourse to religious metaphor - Putnam's "conversion" to anti-realism took place under the tutelage of Michael Dummett. In Chapter Two, I focus upon the exchange between Putnam and Dummett at Jerusalem, and upon the sometimes puzzling paper, "Realism and Reason", in which the "critical" position receives its first mature exposition. I examine the import of Dummett's views for Putnam's arguments against metaphysical realism, especially the "model-theoretic" argument. I also suggest that the exegetical difficulties attending the paper, "Realism and Reason", arise from the essentially transitional nature of that paper. In Chapter Three, I outline Putnam's more considered exploration of these matters in Reason, Truth, and History. and later writings. I discuss the distinction between "external" and "internal" philosophical perspectives, and I offer a preliminary account of the "brain in a vat" argument.

In Part Two, I examine in greater detail the principal weapons in the anti-realist armoury, and assess the effectiveness of realist defences against those weapons. I focus upon three arguments: the "brain in a vat" argument (Chapter Four), the "model-theoretic" argument (Chapter Five), and Dummett's "Manifestation" argument (Chapter.

epistemological infirmities; (C) The secondary literature on the "manifestation" argument reveals serious misreadings of the anti-realist project: in particular, there is a tendency to conflate "semantic realism" and "ontological realism", and to misconstrue the motivation underlying central features of the anti-realist arguments.

I conclude that the arguments of Putnam and Dummett raise serious doubts about the capacity of the realist programme to answer to our cognitive interests as philosophers, and that the burden rests upon the realist to show that we should continue to pursue that programme.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

Introduction

In the introduction to <u>Meaning and the Moral Sciences</u>, Hilary Putnam offers the following characterisation of what he terms "a recent turn in my thinking:"

> Broadly speaking, the thrust is that a "verificationist" view of truth and a correspondence theory of truth (in the sense of empirical realism) are not incompatible. Inasmuch as Kant had a view which included a correspondence view of truth within the empirical realm (on my reading anyway) and a stress on the mind-dependence of all truth, there is, I like to think, something Kantian in the view with which I end up. But it is a demythologised Kantianism, without "things-in themselves" and "transcendental egos".
>
> 1

The "turn" in question, which finds expression in the paper, "Realism and Reason", is a turn <u>away</u> from what, in deference to Putnam's perception of his intellectual heritage, I shall term the "pre-critical" position - the position that he himself describes as "sophisticated realism". Fundamental to the "sophisticated realism" of the pre-critical writings, as we shall see, is an explicit or implicit endorsement of a metaphysically pregnant version of the correspondence theory of truth a view that Putnam terms "metaphysical realism", or, in later writings, "externalism". Part of the "critical" turn in Putnam's thinking, then, consists in the rejection of externalism. But the project also embraces a constructive moment, well defined in the following remarks from a later paper:

> Goodman urges, shockingly, that we give up the notion of "the world"...We make many versions; the standards of rightness that determine what is right and what is wrong are corrigible, relative to task and technique, but not subjective. The question this tendency raises is whether a narrow path can indeed be found between the

swamps of metaphysics and the quicksands of cultural
relativism and historicism. -

Putnam's contention is that the "narrow path" in question is provided
by the view that he terms "internal realism", "internalism", or, in
his most recent treatment of these issues, "pragmatic realism". The
latter is characterised as "at bottom just the insistence that
common-sense realism is not incompatible with conceptual relativity",
where "common-sense realism" is glosses as "a view that takes our
familiar common-sense scheme, as well as our scientific and artistic
and other schemes, at face value, without helping itself to the notion
of the thing 'in itself'."[3] "Pragmatic realism", he maintains, is
the heir to the "Jamesian programme" in philosophy.

I shall have relatively little to say in the present work concer-
ning the "constructive" moment in Putnam's Jamesian programme - the
elaboration of "pragmatic realism" or "internalism". Rather, I shall
take my theme from the programme's "destructive" moment - the rejec-
tion of the traditional realist programme in metaphysics and, deriv-
atively, in the philosophy of language, where "semantic realism"
employs the externalist framework in providing a philosophical account
of the workings of language. To the extent that Putnam's treatment
of these matters is both thematically and historically intertwined
with the work of Michael Dummett, I shall have occasion to examine
the writings of both philosophers. My goals are both exegetical and
critical: I shall attempt both to clarify the arguments brought
against the realist programme, and to assess the force of the various
strategies that realists have adopted in defence of that programme.

Part One, which is primarily exegetical in intent, traces the development of Putnam's thinking on our themes from the pre-critical "sophisticated realism" that culminates in the Locke Lectures, to the internalism developed in Reason, Truth, and History and later works. Chapter One examines Putnam's treatment of "realism" and related notions in the period prior to the 1976 Jerusalem conference on the philosophy of language, where - too close to Damascus to permit one to entirely avoid recourse to religious metaphor - Putnam's "conversion" to anti-realism took place under the tutelage of Michael Dummett. In Chapter Two, I focus upon the exchange between Putnam and Dummett at Jerusalem, and upon the sometimes puzzling paper, "Realism and Reason", in which the "critical" position receives its first mature exposition. I examine the import of Dummett's views for Putnam's arguments against metaphysical realism, especially the "model-theoretic" argument. I also suggest that the exegetical difficulties attending the paper, "Realism and Reason", arise from the essentially transitional nature of that paper. In Chapter Three, I outline Putnam's more considered exploration of these matters in Reason, Truth, and History and later writings. I discuss the distinction between "external" and "internal" philosophical perspectives, and I offer a preliminary account of the "brain in a vat" argument.

In Part Two, I examine in greater detail the principal weapons in the anti-realist armoury, and assess the effectiveness of realist defences against those weapons. I focus upon three arguments: the "brain in a vat" argument (Chapter Four), the "model-theoretic" argument (Chapter Five), and Dummett's "Manifestation" argument (Chapter.

Six). In Chapter Four, I distinguish two kinds of critical response to the "brain in a vat" argument. Firstly, some writers have claimed that the argument resists valid formulation. I attempt to rebut this suggestion by demonstrating confusions in the critical literature, and by offering a valid formulation of the argument. I reject, however, the recent suggestion that the argument can be validly formulated as a dilemma. Secondly, some philosophers have maintained that, even if the argument does admit of a valid formulation, it fails to establish the philosophically significant conclusions that Putnam wishes to establish. I argue that we should take at face value Putnam's expressly modest claims for the argument, and that we should resist the suggestion that the argument is intended either as a refutation of philosophical skepticism, or as a self-contained argument against metaphysical realism.

In Chapter Five, I consider three kinds of realist strategies in response to the "model-theoretic" argument. Firstly, some critics have sought to defend metaphysical realism by offering purported "counter-examples" to the argument's conclusion - the claim that an epistemically "ideal" theory could not be false. Secondly, it has been argued that the argument is technically flawed, either in its appeal to results in model theory or in its employment of the notion of an "ideal" theory. Thirdly, the most common response to the "model-theoretic" argument - what might be termed the "canonical" strategy - is to maintain that Putnam ignores the embedding of language-use in our more general commerce with the world, and that such embedding provides constraints on reference that allow the realist to evade the "model-theoretic" argument. I argue that these strategies, as employed by Putnam's critics,

have tended to involve misrepresentations of both the argument and its intended dialectical role. I conclude, however, that Putnam fails to demonstrate the <u>incoherence</u> of a suitably consistent (or dogmatic?) metaphysical realism. But the epistemological implications of external- ism, as revealed in the debate over the "model-theoretic" argument, suggest that it may be inadequate because it fails to answer to our philosophical interests in seeking an account of language, truth, and reference.

In Chapter Six, I further explore this suggestion by examining in some detail the attack by Dummett and Crispin Wright on "semantic real- ism". I argue that, as with Putnam, much of the secondary literature reveals misunderstandings of the anti-realist programme - in particular, the conflation of different forms of realism, and a misreading of the motivation behind central features of the anti-realist programme. I conclude that the arguments of Putnam, Dummett, and Wright raise serious doubts about the capacity of the realist programme to answer to our cog- nitive interests as philosophers, and that the burden rests on the advoc- ates of realism to show why we should continue to pursue such a pro- gramme, rather than following Putnam in making the "internalist" turn.

Endnotes - Introduction

1. Putnam, 1978, pp. 5-6.

2. Putnam, 1982b, p. 226.

3. Putnam, 1985b, Lecture I.

Chapter One   Sophisticated Realism

I-a   "Realism" Revived

One of the more daunting tasks that confronts the aspiring exegete
of Putnam's _ouevre_ and the secondary·industry.that it has spawned is to
keep track of the various forms of "realism" that circulate, usually
bereft of identifying nametags, in the literature. It might be thought
that the problem originates in Putnam's putting asunder of "metaphysic-
al" and "internal" realism, but a proliferation of realisms can be
found even in the earlier 'pre-critical' writings, and here, as later,
it may contribute towards the reader's sense of puzzlement, of not quite
grasping how everything in the Putnamian picture is supposed to fit.
together. It is readily apparent to the reader of the first two volumes
of Putnam's _Philosophical Papers_, wherein are gathered almost all of his
significant writings prior to the _Locke Lectures_, that one of the unify-
ing threads in the treatment of diverse topics is a concern to defend
and develop something termed "realism". We are told, for example, that
all of the papers in the first volume, _Mathematics, Matter, and Method_.
are written from "a realist perspective,"[1] and that the papers in the
second volume, _Mind, Language, and Reality_, are largely concerned with
the development of "realist" theories of meaning and reference[2]; and,
in a recent lecture, Putnam has characterised the work in which he
engaged in the 1960's as an attempt "to revive and elaborate a kind of
realism."[3]

As soon as we try to clarify our understanding of the nature of
this "realism" and of the reasons why we should endorse such a view,
however, things start to get murky. For example, we are offered the
following account of what realism amounts to with respect to science:

6

> The statements of science are in my view either true
> or false (although it is often the case that we do not
> know which), and their truth or falsity does not con-
> sist in their being highly derived ways of describing
> regularities in human experience. Reality is not a part
> of the human mind; rather, the human mind is a part -
> and a small part at that - of reality.   4

The argument for such realism, we are told, is that we are committed to

the existence of the entities postulated by our scientific theories by

the indispensability of such postulation for the practice of science.

Mathematical realism, for example, is justified as follows: "Quantific-

ation over mathematical entities is indispensable for science, both

formal and physical; therefore we should accept such quantification;

but this commits us to accepting the existence of the mathematical

entities in question."[5] Putnam claims that we should not be moved by

the "fictionalist" counter that the indispensability of certain con-

cepts shows only that they are "useful fictions":

> It is silly to agree that a reason for believing that
> p warrants accepting p in all scientific circumstances
> and then to add "but even so it is not good enough."
> Such a judgment could only be made if one accepted a
> trans-scientific method as superiod to the sceintific
> method.   6

This sort of account of what "realism" is and why we should be

"realists" is not easily reconciled with an alternative answer to these

same questions to be found in Putnam's early paper, "Do True Assertions

Correspond to Reality?"[7] Addressing himself to the question "whether any

sense can be made of the traditional view that a true assertion is one

that corresponds to reality," Putnam argues that realism is "an empir-

ical hypothesis":

> The strategy...is to view the language-speaker as con-
> structing a symbolic representation of his environment.
> He cannot do this unless he causally interacts with the
> environment, and the accuracy or inaccuracy of his rep-
> resentation will affect the viability and success of
> his efforts in dealing with his environment; thus such
> an account of the relation of language-speakers to the
> world is part of a causal model of human behaviour. In
> so far as the assumed correspondence between the rep-
> resentations in the speaker's mind and their external
> referents is part of the model, realism thus becomes an
> empirical hypothesis.   8

It is not difficult to see the connection between "realism", so con-

strued, and so-called "realist" theories of reference which stress the

"contribution of the environment" to fixing the referents of terms. But

how can this sort of realism peacefully cohabit with the essentially

pragmatic brand of "realism" sketched in the previous paragraph? Will

not one who attempts to adhere to both forms of "realism" lay himself

open to the objection that, while quantifying over x's may indeed be

indispensable to practice, this in no way demonstrates that talk of x's

"corresponds to reality"? It is interesting to note that, in his recent

Carus Lectures, Putnam suggests that the only philosophically signific-

ant questions concerning the various forms of ontological realism are

of the pragmatic variety, to be answered on the basis of indispensabil-

ity arguments. But this is not how he would have proposed to resolve

the apparent tension between the two notions of realism in his 'pre-

critical' period. To understand how he would have tried to resolve this

tension, we need to set Putnam's professed realism in its philosophical

context.

The "realism" that Putnam attempted to "revive and elaborate" in

the 1960's and early 1970's was primarily a reaction against certain

doctrines that dominated the American philosophical scene in the pre-

ceding two decades. Foremost among these doctrines was "verificationism",
which provided the underlying theory of meaning for positivist and con-
temporary empiricist views of science. Putnam characterises verificat-
ionism as "the doctrine that to know the meaning of a scientific prop-
osition (or of any proposition, according to most verificationists) is
to know what would be evidence for that proposition"[9]. Given that veri-
ficationist philosophers in the post-war years tended to identify "evi-
dence" with statements about observables, rather than with statements
about sense-experience, the general import of verificationism for
science was that scientific statements about unobservables must be
understood as "highly derived" talk about observables. In arguing
against such a view, Putnam remarks certain perceived failings of veri-
ficationism and advances an alternative, non-verificationist theory of
meaning to serve as the foundation for a realist view of science. He
claims that "naive verificationism" - the view that sentences about
unobservables are identical in meaning to sentences about observable
-evidential conditions - distorts "the character of actual scientific
methodology and inference" by construing the relation between a theo-
retical statement and its evidence as a matter of linguistic conven-
tion, whereas the relation is in fact one of "probabilistic inference
within a theory."[10] The "sophisticated verificationist", on the other
hand, who conceives the relation between theoretical statements and
their evidence correctly and who incorporates this relation into his
theory, is committed to a 'network' account of meaning that obliterates
the distinction between disagreement about the meanings of words and
disagreement about the facts; and in the absence of this distinction,
Putnam maintains, the notion of linguistic meaning becomes vacuous.[11]

In place of verificationism, Putnam proposes a theory of meaning according to which we should "see meanings as largely determined by reference, and reference as largely determined by causal connections."[12] The "causal" theory of reference, developed independently by Putnam and Saul Kripke[13], stresses both the "linguistic division of labour" - the social features of the process whereby reference is fixed - and "the contribution of the environment" - the respects in which the extension of certain kinds of terms is "in part, fixed by the world":

> There are objective laws obeyed by multiple sclerosis,
> by gold, by horses, by electricity, and what it is rat-
> ional to include in these classes will depend on what
> those laws turn out to be. It is because we do not know
> these laws exactly that we have to leave the extension of
> these classes somewhat open, rather than fixing it ex-
> actly by making the terms synonymous with sets of nec-
> essary and sufficient conditions.   14

Any "natural kind" term has associated with it certain "paradigms", generally recognised exemplars of the kind, and something belongs to the kind if it possesses the "same nature" as (at least the majority of) these paradigms, where "same nature" is to be cashed out in terms of obedience to the same objective laws. The meaning of a term, accor- ding to Putnam, comprises both the reference and what he terms the "stereotype", a collection of markers which competent speakers associ- ate with the term.[15] Given such an account of meaning, we can formulate a "realist" interpretation of theoretical discourse in science. State- ments about unobservables, e.g. electrons, are not "highly derived" talk about instrument readings or other observables; rather, they are talk about those entities nomically related to certain observable phen- omena - those entities that are "causally connected" in the appropriate

way to our use of the term 'electron', in the present case.

This account of the dialectic between scientific realism and verificationism may profitably be compared with Putnam's argument for mathematical realism. The empiricist 'foe', here, is not verificationism, but "conventionalism", the view that the principles of mathematics are "conventions or rules of language and therefore say nothing."[16] Putnam presents two objections against such a view. Firstly, he maintains that only some mathematical sentences could be immediately true by convention, since we can enact only a finite number of meaning-stipulations whereas the number of mathematical sentences is infinite. The residual mathematical sentences must therefore be implied by the sentences immediately true by convention. But the notion of implication required in this context can be defined only by means of set theory, and thus the conventionalist thesis reduces to the uninteresting claim that mathematical sentences are true in virtue of immediate convention and mathematics.[17] Secondly, Putnam echoes Quine in rejecting the idea that any sentences are "true by convention" in the empiricist's sense - "absolutely a priori" in that it would never be rational to give them up, or even to doubt them.[18] This is not to deny that "there is a distinction between truths which are a priori relative to a particular body of knowledge and truths which are empirical relative to a particular body of knowledge", nor is it to deny that mathematics is much more contextually a priori than physics. But to be a priori in this sense is not to lack empirical content, according to Putnam: mathematics resembles physics in that it involves "an interplay of postulation, quasi-empirical testing, and conceptual revolution leading to the formation of contextually a priori paradigms."[19]

We are now in a position to grasp more clearly the role of the "indispensability" argument in establishing various forms of realism. The indispensability argument may be regarded as a general licence to assert the existence of any entities quantification over which occurs within scientific (and maybe also non-scientific) practice and is taken to be indispensable for that practice. This general licence is subject to certain defeasibility conditions, however, namely, that it does not apply where talk about the entities in question requires "philosophical reinterpretation" and cannot be taken "at face value", either because we are really talking about observables (verificationism), or because we are really talking about linguistic conventions, or because we are merely employing useful fictions. To show that none of these defeasibility conditions can be properly invoked with respect to a given form of discourse is to show that the indispensability argument can be applied to yield realism concerning the entities posited within that discourse.

But what is it to take the statements in a given form of discourse "at face value"? This question leads us to a further sense of "realism" at play in Putnam's 'pre-critical' writings. In the Introduction to the first volume of his Philosophical Papers, Putnam suggests that the unifying theme in the contained papers, which are all written from "a realist perspective," may be "the view that science, including mathematics, is a unified story, and that story is not a myth but an approximation to the truth."[20] But to characterise "realism" in terms of truth is unsatisfactory unless we are also informed as to how the latter notion is to be understood. For, as Putnam points out in the Introduction to Meaning and the Moral Sciences (henceforth MMS),

truth itself admits of so-called "realist" and "non-realist" interpret-
ations. On a "realist" interpretation, truth is "some kind of correspon-
dence between ideas and 'what is the case'", whereas, on a "non-realist"
or "verificationist" interpretation, truth is "radically mind-depend-
ent...as, for example, what would be verified under ideal conditions of
inquiry."[21] Putnam claims that recent empiricists, while espousing a
verificationist theory of meaning, have not endorsed a verificationist
theory of truth. Rather, they have tended to adopt a "deflationary"
view of truth whereby, even if lip-service is paid to the notion of
truth as correspondence, truth is taken to be a philosophically uninter-
esting notion, requiring only the satisfaction of Tarski's "equivalence
condition". Tarski states[22] that an adequate "truth-defintion" for a
(formalised) language must meet the requirement that, for any sentence
in the object-language, it is provable in the meta-language that "'S'
is true iff S". On a deflationary account, the notion of truth may be
viewed as a device allowing semantic ascent from the object-language to
the meta-language, but it does no "philosophical work" in so far as a
sentence's satisfying the "equivalence condition" "tells us nothing
about how [that sentence] is used or understood; nothing about its
assertibility, etc."[23] Putnam's contention that "realism is an empirical
hypothesis" is part of "a realist reaction" to the empiricists' accep-
tance of such a deflationary view of truth. The realist maintains that
Tarski has given "a perfectly correct account of the formal logic of
the concept 'true'", but holds that "Tarski's work requires supplement-
ation; and when this philosophical supplementation is provided, we see
that the notion of truth is not philosophically neutral, and that a
correspondence account is needed to understand how language works and

how science works."[24] "Realism as an empirical hypothesis" is an
attempt to refurbish and render philosophically respectable the tradit-
ional realist conception of truth as "correspondence". And it is this
refurbished "correspondence" theory of truth that Putnam labels "soph-
isticated realism".

Sophisticated realism receives its fullest exposition in the
Locke Lectures, and, more specifically, in those sections of the second
and third lectures originally presented, in slightly abridged form,
in the lecture "What is Realism?".[25] Before I turn to the Locke Lectures,
however, let me make a number of brief parting observations on Putnam's
pre-1976 professions of "realism". Firstly, Putnam stresses that he is
not advocating any form of materialism: a consistent realist, he main-
tains, must embrace in his ontology such things as mathematical objects,
fields (in the physicist's sense), and physical magnitudes.[26] In the
same vein, he characterises his championing of functionalism in psych-
ology as a rejection of reductionism, and urges that we see psychology
as having "an autonomous explanatory function."[27] Secondly, Putnam
opposes not only materialism and reductionism within science and mathe-
matics, but also the idea that all knowledge is, or is reducible to,
scientific knowledge: "The idea that the concepts of truth, falsity,
explanation, and even understanding are all concepts that belong ex-
clusively to science seems to me to be a perversion."[28] He suggests
that there can be knowledge of ethical principles, where objectivity
"is connected with such things as width of appeal, ability to with-
stand certain kinds of rational criticism..., feasibility, ideality,
and, of course, with how it actually feels to live by them or attempt
to live by them;"[29] and that literature gives us "conceptual knowledge..

knowledge of a possibility."[30] But, as Putnam acknowledges, these sugg-
estions are programmatic at this stage in his writing, although the
notion of ethical objectivity·has been accorded considerable attention
in Putnam's attempts at developing the "Internalist" perspective (see
chapter three below).

Thirdly, in trying to clarify the relationship between "indispen-
sability" arguments, as part of the case for various brands of "ontol-
ogical realism", and the thesis that realism is "an empirical hypo-
thesis", as part of the case for a realist theory of truth, I have per-
haps oversimplified Putnam's pre-critical position. The point is not
that the distinction in question is not a coherent one - on this point,
see further chapter six below - but, rather, that Putnam's use of
"indispensability" arguments may not (or may not always) be pragmatic
as suggested above. Employed pragmatically, the "indispensability"
argument reasons from our practical need to posit certain entities to
the reality of those entities on the grounds that one cannot cogently
draw a philosophical distinction between "reality" as defined through
human practice and "reality" as it "really" is. But one may also employ
the "indispensability" argument non-pragmatically by arguing that our
need to posit certain entities is evidence for the reality of those
entities - that the entities really exist is part of the "best explan-
ation" of the relevant features of our practice. It is this use of
"indispensability" arguments that we seem to find in Putnam's argument
for "realism as an empirical hypothesis", quoted above, and it is a
similar use of such considerations that we shall also find in the
"miracle" argument set forth in the second Locke Lecture.

I-b   The Locke Lectures

In the Introduction to MMS, Putnam states that the Locke Lectures were originally conceived as a response to Hartry Field's article, "Tarski's Theory of Truth"[31], wherein Field argues that Tarski's work is unsatisfactory from a realist point of view because it fails to provide a "physicalistic" account of primitive reference. In responding to Field in the Locke Lectures, Putnam expands upon themes examined in the preceding section - the credentials of "deflationary" views of truth, and the reconceptualisation of the correspondence theory of truth as an "empirical hypothesis". In arguing that the realist should resist Field's demand for a "physicalistic" theory of reference, Putnam is led to address other issues of contemporary philosophical interest, such as the tendency towards "scientism" in the social sciences and the nature of radical translation. Although the individual arguments in the six Lectures are reasonably clear, it is not always easy to retain a grasp upon the overall direction of the discussion. In particular: (a) it is unclear precisely how Putnam proposes to answer a crucial question twice posed in Lecture II, namely, "What is the relation between realist explanations of the scientific method, its success, its convergence, and the realist view of truth?"; (b) it is unclear, as Michael Devitt has pointed out[32], whether Putnam believes he has an answer to the question posed at the beginning of Lecture III, namely, "What is it to understand truth and the logical connectives realistically?"; and (c) it is unclear how the discussion of the final three-and-a-half Lectures bears upon the issues raised in the prior Lectures. In the brief summary of the argument of the Locke Lectures that follows, I shall attempt to illuminate these matters. In the interests of greater

transparency of exposition, I shall to some extent abstract from the
original order of presentation.

In the first Lecture, Putnam outlines the relevant features of
Tarski's work, Field's criticism of Tarski, and a subsequent response
to Field by Stephen Leeds. Putnam offers the following characterisation
of Tarski's work:

> What Tarski does is show how, in the context of a for-
> malised language, one can define 'true' (or a predicate
> which can be used in place of 'true') using only the
> notions of the object language and notions of pure
> mathematics. In particular, no semantical notion - no
> such notion as 'designates', or 'stands for', or
> 'refers to' - is taken as primitive by Tarski (although
> 'refers to' gets defined - defined in terms of non-
> semantical notions - in the course of his work).   . 33

Field's criticism focusses on Tarski's proposed definition of 'primit-
ive reference'. Given a list of the primitive predicates of a partic-
ular formalised language, $\emptyset_1$, $\emptyset_2$,...., $\emptyset_n$, 'primitive reference' is
defined for that language by giving a list with the following struc-
ture: P primitively refers to x iff (1) P is the phrase '$\emptyset_1$' and $\emptyset_1 x$,
or (2) P is the phrase '$\emptyset_2$' and $\emptyset_2 x$, or...., or (n) P is the phrase
'$\emptyset_n$' and $\emptyset_n x$. Field contends that the same scientific standards that
apply in the natural sciences should apply in the theory of language,
and that, given those standards, Tarski's "definition" of 'primitive
reference' is no more acceptable than would be a "valence definition"
of the following form: "...$(\forall E)(\forall n)(E$ has valence n = . E is potassium
and n is +1, or...E is sulphur and n is 2) where in the blanks go a
list of similar clauses, one for each element."[34] Field's claim is
that a realist can be satisfied with nothing less than

a "physicalistic" theory of reference, i.e. a theory that
says that a speaker refers to something when he uses a
term just in case his use of that term stands in a defin-
ite causal relation (to be defined by empirical science,
in just the spirit that empirical science has 'defined'
<u>water</u> as H_2O and <u>temperature</u> as mean molecular kinetic
energy) to that something, or to things of the same kind
as that something (in a sense of 'same kind' which would
also have to be defined.   35

Putnam cites, as a possible line of response to Field, an unpub-
lished paper by Stephen Leeds. Leeds argues that notions like truth and
reference do not serve the same <u>purpose</u> as notions like valence, and
that, since our reasons for rejecting the proposed definition of 'val-
ence' relate to the purposes which the notion of valence serves, these
reasons need not, and do not, apply to the proposed definition of 'prim-
itive reference':

> Valence is an <u>explanatory</u> notion (i.e. a <u>causal-explanat-</u>
> <u>ory</u> notion). Since we intend the existence of various
> valences to figure in chemistry as a <u>cause</u>, we have to say
> what valence <u>is</u>, not just give the numerical values. But
> Leeds is <u>denying</u> that <u>reference</u> is a causal-explanatory
> notion. We need notions like truth and reference to <u>ex-</u>
> <u>press</u> certain things (which could in principle be express-
> ed in other ways - by using infinitary languages). For
> <u>this</u> purpose, it is immaterial if primitive reference is
> defined in what Field calls a "crazy" way. Reference
> isn't (or, anyway, Field hasn't shown that it is) a
> causal-explanatory notion.   36

Putnam wants to steer a course between Field and Leeds. On the one hand,
he thinks that Leeds is right in claiming that the purposes for which
we need notions of truth and reference allow us to give "a 'transcen-
dental argument' for Tarski's procedure" without compromising realism.
On the other hand, he believes that "it is not <u>wholly</u> wrong to think of
reference as a causal-explanatory notion."[37] The element of truth in
this "realist intuition" is captured, so he argues, in the "sophistic-

ated realist account of truth mentioned in the previous section - the

correspondence theory of truth construed as, in some sense, "an empir-

ical hypothesis". But, so Putnam also wishes to argue, the "sophistic-

ated realist" is <u>not</u> committed to providing a physicalistic account of

reference - indeed, the "sophisticated realist" will understand precise-

ly what is wrong in the demand for such an account.

We may begin by elaborating upon Putnam's presentation of "sophis-

ticated realism", a presentation whose circuitous nature has already

been remarked. Realists, he maintains, typically proclaim their belief

in a correspondence theory of truth, and typically <u>argue</u> <u>for</u> realism

by arguing <u>against</u> some form of idealism. As he points out, the "real-

ism" at issue here is what is customarily termed "scientific realism",

and the anti-idealist usually maintains that only the realist can do

justice to the <u>success</u> of science:

> The modern positivist has to leave it without explanat-
> ion (the realist charges) that 'electron calculi' and
> 'space-time calculi' and 'DNA calculi' correctly predict
> observable phenomena if, in reality, there are no elec-
> trons, no curved space-time, and no DNA molecules. If
> there are such things, then a natural explanation of the
> success of these theories is that they are <u>partially</u>
> <u>true accounts</u> of how they behave...But if these objects
> don't really exist at all, then it is a <u>miracle</u> that a
> theory which speaks of gravitational action at a dis-
> tance successfully predicts phenomena; it is a <u>miracle</u>
> that a theory which speaks of curved space-time suc-
> essfully predicts phenomena; and the fact that the laws
> of the former theory are derivable 'in the limit' from
> the laws of the latter theory has no methodological
> significance. [38]

If realism is an <u>explanation</u> of the success of science, then "realism

must itself be an over-arching scientific hypothesis"..But then it is

unclear how realism bears upon "the correspondence theory of truth...

or _any_ theory of truth for that matter."[39]

Putnam attempts to clarify the relationship between "scientific realism", as an over-arching scientific hypothesis, and a realist notion of _truth_ by bringing into sharper focus what each of these forms of realism involves. Firstly, he appeals to the work of Richard Boyd to illuminate the nature of scientific realism. Boyd claims that scientists try to preserve not only the observational 'successes' of earlier theories, but also "the approximate truth of the theoretical laws of the earlier theories in certain circumstances." Furthermore, Boyd maintains, this strategy has worked in that it has led to significant scientific discoveries:

> Boyd tries to spell out realism as an over-arching
> empirical hypothesis by means of two principles:
> (1) Terms in a mature science typically _refer_.
> (2) The laws of a theory belonging to a mature science
> are typically approximately _true_.
> What he attempts to show in his essay is that scientists
> act as they do because they _believe_ (1) and (2) and
> that their strategy works because (1) and (2) are _true_. [40]

Boyd makes out the case for _convergence_ in science, and he also provides a "causal-explanatory" role for the notions of truth and reference since these notions occur in (1) and (2) which are themselves "premisses in an _explanation_ of the behaviour of scientists and the success of science."

Putnam then inquires as to the notions of truth and reference that would be compatible with the _denial_ of convergence. He argues, against such writers as Kuhn and Feyerabend, that our present situation in science supports the hypothesis of convergence, since our present theories assert the existence of entities similar in "role" to the entities

postulated under the same names by earlier generations of scientists.
This, according to Putnam, allows us to extend the principle of "the
benefit of the doubt" to our scientific forebears and see them as re-
ferring to the same entities as ourselves; albeit that their theories
concerning the natures of those entities are now discredited. But if
events were to transpire such that successive generations of scientists
failed to assert the existence of entities similar in "role" to the
entities postulated by their immediate forebears, then extending the
benefit of the doubt would become unreasonable, and we could no longer
claim convergence. In these circumstances, we would have to abandon
"realism", construed as "the belief in any describable world of unob-
servable things," and accept in its place "the belief that all the
'unobservable things' (and, possibly, the observable things as well)
spoken of in any generation's scientific theories, including our own,
are mere theoretical conveniences, destined to be replaced and supplan-
ted by quite different and unrelated theoretical constructions in the
future."[41]

Putnam maintains, however, that the effect of thus abandoning
"realism"

> would not be a total scrapping of the predicates true
> and refers in their formal aspects. We could...keep
> formal semantics (including Tarski-type truth-defini-
> tions); even keep classical logic; yet shift our notion
> of 'truth' over to something approximating 'warranted
> assertibility'.
> [42]

But if retaining the 'formal' aspects of truth is quite compatible with
a "quasi-intuitionalist" understanding of that notion, then, conversely,
a "realistic" understanding of truth cannot consist in the mere adher-

ence to these 'formal' aspects. The crucial question, therefore, is what it is that demonstrates that one understands the notion of truth "realistically". Putnam claims that the mark of a "realist" understanding of truth is

> one's acceptance of such statements as:
> (A) Venus might not have carbon dioxide in its atmosphere even though it follows from our theory that Venus has carbon dioxide in its atmosphere.
> and
> (B) A statement can be false even though it follows from our theory (or from our theory plus the set of true observation sentences.)   43

We may now be in a better position to grasp the connection which Putnam is positing between "scientific realism" and a "realist" theory of truth. Firstly, Putnam makes two claims concerning a "realist" understanding of truth which consists in the acceptance of statements such as (A) and (B): (a) our present concept of truth is "realist" in this sense, for we believe that there is a greater than zero probability that our current theories might be false; and (b) our reason for accepting statements such as (A) is that they are themselves part of our overall theory in so far as the latter incorporates "knowledge as part of the subject of our knowledge":

> We view knowledge itself as the product of certain types of causal interactions, at least in such cases as "Venus has carbon dioxide in its atmosphere." And it follows from our theory of the interaction whereby we learned this fact - for example, the standard causal account of perception and the theory of error - that we might have, for any number of reasons, made up a theory from which it follows that Venus has carbon dioxide in its atmosphere, even if it did not. In short, (A) is itself a "scientific" (or even a commonsense) fact about the world (albeit a modal fact about the world).   44

The claim, therefore, is that "science taken at 'face value' implies realism;" the bearing of Boyd's argument from scientific success to scientific convergence upon the tenability of a "realist" notion of truth is that we require convergence if we are to ward off arguments against taking science at face value.

This emerges clearly in Putnam's discussion of three possible strategies that one might adopt in an attempt to avoid accepting the "realist" notion of truth to which our overall theory of the world would seem to commit us. Firstly, one might appeal to some general programme for the philosophical reinterpretation of the theoretical vocabulary of science - for example, phenomenalism or logical empiricism; but all such programmes have proved unsuccessful, Putnam maintains. Secondly, one might follow Peirce in accepting statements like (A) and (B) while at the same time identifying truth with warranted assertibility in the ideal limit of scientific inquiry. Putnam responds to this suggestion by citing certain difficulties that the Peircean account, taken to be a non-realist account of truth and falsity, purportedly faces: (i) we seem to require a "realist" notion of truth in terms of which to specify just what is to count as the "ideal limit" of inquiry, and (ii) the Peircean account, like "realism", requires convergence. But even if we grant Putnam these points, this fails to answer Devitt's complaint, noted above, that, if (A) and (B) can be accepted by one who espouses a non-realist view of truth, then acceptance of (A) and (B) cannot be, as Putnam seems to be claiming, the mark of a "realist" understanding of truth. Thus the question posed at the beginning of Lecture III - "What is the mark of a "realist" understanding of truth?" - is not answered in the sequel. (We may note, in anticipation, that

Putnam in a sense resolves this difficulty - without acknowledging that it _is_ a difficulty - in the paper "Realism and Reason", where the 'modal' characterisation of "realism" is retained but where Peircean theories of truth are reclassified as _realist_!)

It is the third strategy that reveals the link between "realist" truth and the arguments for "scientific realism". The latter, as we have seen, involve an inference, from the success of scientific practice in certain specific respects, to the conclusion that mature sciences exhib- it _convergence_. But if Kuhnian or Feyerabendian arguments _against_ con- vergence were to succeed, and we were unable to preserve the reference of theoretical terms across scientific "revolutions", then "eventually the following meta-induction becomes overwhelmingly compelling: _just as no term used in the science of more than fifty_ (or whatever) _years ago referred, so it will turn out that no terms used now_ (except maybe observation terms, if there are such) _refer._"[45] Since the realist notion of reference (and thus of truth) leads to such disastrous con- sequences when combined with the assumption that there is no convergence, and since no theory of reference having such consequences is acceptable, our only alternative, in the envisaged circumstances, would be to give up "realist" truth: "We must fall back on an intuitionist or quasi- intuitionist reading of the logical connectives, which would save the bulk of extensional scientific theory, and save the formal part of our theories of reference and truth, at the cost of giving up (A) and (B)."[46] It is Boyd's argument for a sophisticated notion of convergence, and, thereby, for "scientific realism", that defeats this kind of anti- realist argument, according to Putnam. If we accept the case for conver- gence, then we can take science "at face value"; and, as we have seen,

Putnam's contention is that science, taken at face value, implies a
"realist" understanding of truth.

Before returning to Putnam's discussion of Field, a further prom-
inent feature of the "sophisticated realism" expounded in the Locke
Lectures deserves mention. The "sophisticated realist", as we have seen,
espouses a "realist" account of truth because such an account is implied
by "our total theory of knowledge", where the latter incorporates "our
theory of nature and of our interactions with nature." For similar
reasons, so Putnam argues, the "sophisticated realist" must recognise
certain limits upon his or her realist commitments:

> A twentieth-century realist cannot ignore the existence
> of equivalent descriptions: realism is not committed to
> there being one true theory (and only one)...Assuming
> there is a "fact of the matter" as to "which is true"
> (if either) whenever we have two "intuitively different"
> theories is naive. "Theories" which differ on which
> pairs of events are simultaneous are certainly "intuit-
> ively different", but after Einstein we know that such
> "theories" may, none the less, be equivalent...A sophis-
> ticated realist recognises the existence of equivalent·
> descriptions, because it follows from his theory of the
> world that there are these various descriptions, as it
> follows from a geographer's description of the earth
> that there are alternative mappings (mercator, polar,
> etc.)...
> To show that T_1, T_2, are equivalent descriptions (in the
> same vocabulary) one commonly shows that they agree on
> certain "invariants" and argues (in terms of the explan-
> atory role of the invariants) that this is all they
> should agree on: that the invariants give a "complete"
> picture of what is going on.   47

In an earlier context, Putnam attributes this doctrine of "invariants"
and "equivalent descriptions" to Reichenbach, and clarifies its place
within a realist framework. If A and B are equivalent descriptions in
the same vocabulary, then questions on which A and B differ, including
ontological questions, will have "no representation-independent answer."

Whether two accounts of a given domain are properly regarded as equiv-
alent descriptions will depend upon the theories we accept and the
notion of 'invariant' that those theories yield. This does not, however,
entail that "representation-independence", and thus ontology, is a
purely theory-relative matter in any "relativist" sense. Rather, so
Putnam maintains, the correct set of invariants is that which is yield-
ed by "the correct physical theory of the world."[48]

  . How does this bear upon the original issue of Field's "realist"
critique of Tarski? Putnam's answer to Field comes in two stages. He
argues, firstly, that the realist can accept Tarski's definition of.
'truth' on the basis of the sorts of considerations advanced by Leeds.
He then counters the suggestion that the "causal-explanatory" role of
the notions of truth and reference, as that role is conceived by the
"sophisticated realist", commits the latter to the need for a physical-
istic account of reference of the sort demanded by Field. The first
stage of this answer begins with the contention that the realist's
correspondence theory of truth is not a proposed definition of 'truth'
because "there is only one way anyone knows how to define 'true' and
that is Tarski's way". The important thing about Tarski's definition is
that it captures "the formal properties we want the notions of truth
and reference to have," given that "we need for a variety of purposes
to have a predicate in our meta-language that satisfies precisely the
Criterion S" (where the "Criterion S" is the natural generalisation of
Tarski's Criterion T to formulae containing free variables). It is
these requirements that are reflected in Tarski's definition of "prim-
itive reference", for the clauses in the list that the definition com-
prises are structured to conform to Criterion S. Thus Tarski's defin-

ition captures the properties that we require in our notion of truth, whether that notion be understood "realistically" or otherwise, and the definition is therefore perfectly acceptable to the realist.[49]

Putnam then considers the following response open to Field. Might it not be said that Boyd's argument for realism, which assigns a certain causal-explanatory role to the notions of truth and reference, undermines Leeds' claim that these notions are not properly analogous to the notion of valence? And, if so, would not this causal-explanatory role accorded to the notions of truth and reference justify "looking for a physicalistic account of what truth and reference are just as the causal-explanatory role of 'valence' justifies looking for a physicalistic account of what valence is?"[50]

In keeping with methodological tendencies in the Locke Lectures that we have already had cause to remark, Putnam approaches these questions obliquely. Rather than directly address the issue of reference, he embarks upon an extended discussion of the enterprise of translation. Only after arriving at certain conclusions concerning translation does he argue that analogous conclusions are warranted in the case of reference and that this provides an answer to Field. The crucial claim as far as translation is concerned is that the correctness of a translation is a function of its explanatory power:

> What we do in "translation" is to construct a global theory which gives reasonable explanations of the speaker's behaviour in the light of his beliefs (as determined by the translation manual which is one component of the global theory) and his desires and intentions (as determined by the psychological theory which is the other component)...I am not just contending that it is good methodology in finding out what a speaker "means" to try to rationalize his behaviour in this way! I am suggesting that what it is to be a correct translation or truth-

> definition is to be the translation or truth-definition
> that best explains the behaviour of the speaker. ·   51

Putnam then argues that "explanation is an interest-relative notion",
and that this allows us to view in a more plausible light Quine's
"Indeterminacy Thesis". Given certain cognitive interests underlying
the enterprise of translation, there may be a unique translation-manual
that is "correct" in the above sense. But, to the extent that we can
conceive of other cultures whose cognitive interests in translation
differ from our own, we can also entertain the possibility of a plural-
ity of equally correct translation-manuals, as Quine himself maintains.
Invoking the doctrine of "invariants" and "equivalent descriptions",
Putnam counters the suggestion that the native speaker's own interests
and beliefs can determine which of these "equally correct" translation-
manuals is really correct: he argues that the appropriate notion of
"invariant" for psychology is at the level of functional organisation,
and that alternative psychological descriptions at the "intentional"
level are properly viewed as "equivalent descriptions" relative to such
invariants.

Field's thesis is reengaged when the suggestion is entertained
that we should try to render the notion of correct translation scien-
tifically precise. Since translation takes place between diverse human
cultures, giving us the phenomenon of "the universal intercommunicabil-
ity of human cultures," and since translation is a teachable skill, we
seem forced to conclude that "human interests, human saliences, human
cognitive processes, must have a structure which is heavily determined
by innate or constitutional factors."[52] This raises the possibility that
we might overcome the interest-relativity of translation by "'fixing'

the interests - this time taking the <u>invariant</u>, biologically specified interests, saliences, priorities, etc." Putnam rejects such a proposal on the grounds that it would tell us nothing about the <u>nature</u> of translation, that it would be inadequate in the way that all operationalist analyses are, and that it would be "wholly unworkable in practice" It is the latter point that receives the most attention. The problem, Putnam maintains, is that "to simulate translation in general, radical translation, one would have to have a simulation of full human intelligence, saliences, etc." since "<u>everything we know</u> is used in translation."[53] To make explicit all of the constraints on correct translation would require "<u>a detailed explanatory model of the functional organisation of a whole human being</u>," and it is "utopian" to anticipate the development of such a model since (a) we face not merely complexity but "highly structured complexity", (b) if we possessed such a model, this would itself modify our natures, and (c) there is reason to doubt whether such a model would even be <u>intelligible</u> to human beings.[54] The moral to be drawn from all of this, according to Putnam, is that "<u>the notion of translation...cannot be made scientifically precise</u> (at least not in the foreseeable future; and very likely <u>never</u>). <u>Yet it is perfectly usable in daily life and even in parts of the science of linguistics</u>."[55]

Putnam maintains that the preceding conclusions concerning translation apply, <u>mutatis mutandis</u>, to reference. In the relevant respects, Field's programme of "<u>defining</u> reference in terms of 'causal chains'" is just like the programme of "<u>defining</u> translation by a set of 'constraints'":

> As soon as one tries to broaden the causal theory so as
> to cover, say, theoretical terms in science ('electron'),
> then the principle of the benefit of the doubt comes in,
> in some version or other - the principle that says that
> to find out what Bohr referred to by 'electron' in 1904
> we must see what would be reasonable reformulations of
> the descriptions he gave which failed to refer...; and
> giving any precise analysis of the notion of a reason-
> able reformulation of a definite description is, if
> anything, more hopeless than giving a precise list of
> constraints on translation. And the problems are very
> similar: both reference assignment and translation
> depend on choosing "reasonably" to pair up not-exactly-
> matching sets of beliefs. To simulate (or even precise-
> ly to define) "reasonableness" is to simulate (or at
> least define) full human capacity. In short, Field's
> program is a species of scientific utopianism.   56

But to reject Field's programme is not to reject the "causal theory" of

reference, Putnam maintains. For the latter is not a theory of what

reference is, but of "how reference is specified", that is, how speak-

ers pick out the referents of their terms.

We may note, in anticipation of future skirmishes between Putnam

and proponents of physicalist theories of reference (see chapter five

below), that this argument is not without its problems. For suppose that

we grant that our practices of translation and of assigning referents

do indeed involve certain human abilities that we cannot reasonably hope

to model to any degree of scientific precision. In the case of trans-

lation, the futility of anticipating a scientifically precise account

of correct translation follows from what we have just granted because,

so it is argued, a "correct" translation just is the translation that

we accept in so far as we properly employ the human abilities in ques-

tion. But must we also accept that there is no "fact of the matter",

independently of our practice of assigning referents, as to what a

speaker refers to by his use of a given term? The physicalist will

surely maintain that there _is_ such an independent fact of the matter,
consisting in some physicalistically describable relation between the
relevant features of language use and certain items in the world, and
that our practices of assigning referents, "reasonable reformulations"
included, are attempts to determine what the practice-independent facts
of the matter about reference are. Or, approaching the issue from
a different perspective, if there are no such practice-independent facts
of the matter as to the referents of terms — if the true referent of a
particular term as used by a speaker or a linguistic community just is
whatever it would be "reasonable" to assign as a referent - then how
can reference perform the sort of "causal-explanatory" role allotted
to it by Boyd? And if there _are_ such practice-independent facts, why
is it "utopian" to seek a physicalistic account of reference, since it
is the practice of reference assignment that involves "full human cap-
acity"? These issues will resonate in the sequel.

I-c   "Reference and Understanding"

.I shall conclude my brief synopsis of Putnam's 'pre-critical'
writings by summarising the relevant features of the argument of the
paper "Reference and Understanding", delivered at the 1976 Jerusalem
Conference on the Philosophy of Language with a Commentary by Michael
Dummett. The commentator could not have been more aptly chosen, given
Putnam's characterisation of the paper as "in large part a reaction to
[Dummett's] William James Lectures" which he had either heard or read
in typescript.[57] The feature of Dummett's Lectures that attracted
Putnam's attention was the former's assault upon a philosophical pos-
ition that he termed "realism". I examine Dummett's views in some
detail in chapter six below. For the present, we need only note that
Dummett's "realism" in so far as it concerns Putnam is a semantic
thesis. The "realist" adheres to a "truth-conditional semantics",
whereby individual sentences of a language are associated with truth-con-
ditions   that are not epistemically constrained - that is, truth-
conditions that can transcend our capacity to recognise that they obtain.
Furthermore, the Dummettian "realist" holds that a speaker's mastery of
a language is to be represented as a knowledge of such truth-conditions
for the individual sentences of the language.

In "Reference and Understanding", Putnam contends that one can
recognise the force of Dummett's arguments against "realism" in Dumm-
ett's sense - adherence to a truth-conditional semantics - without
thereby compromising "realism" in Putnam's sense - acceptance of an
"over-arching empirical hypothesis" to be understood in terms of a
"realist" notion of truth. In fact, so Putnam maintains, rejecting
Dummett's "realism" may even help the case for his own "realism", since

it enables one to circumvent a traditional objection to the correspond-
ence theory of truth. He maintains that a "Putnamian" realist should
not represent linguistic understanding by ascribing to competent
speakers a knowledge of truth-conditions. Rather, understanding a lan-
guage should be viewed as "the possession of a rational activity of
language use - an activity involving 'language entry rules' (proced-
ures for subjecting some sentences to stimulus control), procedures
for deductive and inductive inference, and 'language exit rules' (pro-
cedures for, e.g. 'taking something in one's hand when one decides it
would be optimal to take this in one's hand')"[59].   Such a model of
linguistic understanding is, he claims, quite compatible with the feat-
ures of language stressed in "The Meaning of 'Meaning'" - the linguistic
division of labour and the importance of 'stereotypes'. Furthermore, it
is essentially "holistic" in nature:

> The conditions under which any particular sentence will
> be uttered and the behaviour that will result if any
> given sentence is uttered does not depend upon any iso-
> lated thing that could be called the 'sense' of the sen-
> tence, but on the total system. Changing the inductive
> logic, or the deductive logic, or the utility function
> will to some extent affect the utterance conditions for
> and the behaviour responses to every sentence of the lan-
> guage (although ordinary statements about macro-obser-
> vables will, of course, have canonical 'verifying' and
> 'falsifying' experiences associated with them).   60

If we accept this "use" model of linguistic understanding, then,
Putnam contends, we can answer the charge that the proponent of the
correspondence theory of truth is both (a) committed to the thesis that
understanding any statement presupposes knowledge of the nature of the
purported "correspondence" relation between language and reality, and
(b) unable to explain in what such knowledge consists, given that it

cannot "consist in the acceptance of any statements, because it is prior to the understanding of all statements." The correspondence theorist is open to such a charge only if he also espouses a "truth-conditional" account of understanding, Putnam maintains. If we construe understanding as a matter of "ability to use" in the matter sketched above, however, there is no problem: "One does not need to _know_ that there is a correspondence between words and extra-linguistic entities to learn one's language...After one has learned one's language one can talk about anything --including the correspondence in question."[61]

But how does the notion of "correspondence" get into the picture here? At this point, Putnam recurs to the thesis that realism is an over-arching empirical hypothesis. Although a "correspondence" between language and reality is not mentioned in the "use" account of linguistic understanding, "it doesn't follow that such a correspondence doesn't exist." In fact, he claims, realist notions of truth and reference may play a crucial part in _explaining_ the place of language in the overall economy of our behaviour: "The _success_ of the 'language-using program' may well depend on the existence of a suitable correspondence between the words of a language and states of affairs.,, It _is_ essential to view our theories as a kind of 'map' of the world, realists contend, if we are to explain how they help us to guide our conduct as they do." A realist notion of truth plays a role in explaining "the _contribution_ of our linguistic behaviour to the success of our total behaviour."[62] _Every_ metaphysical position will offer the same sort of account of such matters. Success in the achievement of our goals is to be accounted for in terms of (1) our generally acting so as to get what we want _if_ our beliefs are true, and (2) the actual truth of many of our beliefs.

Although every theory of truth will attempt to make good on such an account, an adequate theory must also be able to account for "the reliability of our learning," since the latter is entailed by (2) together with the fact that our stock of beliefs is constantly changing. Putnam then sketches how the realist can explain the reliability of our learning "from within our total conceptual system (causal theory of perception and language plus semantic theory of truth), as he reconstructs that conceptual system":

> The role played in this by the idea of a correspondence between language items and extra-linguistic reality is not hard to see. The causal theory of reliability tells us that, for example, when a certain state of affairs obtains (the rug being green) the speaker utters a certain sentence ("the rug is green"). The semantic theory of truth tells us that the sentence is true just in case that state of affairs obtains - the correspondence involved in the causal story is exactly the correspondence set up by the truth definition. [63]

Putnam considers a possible response to the foregoing. It might be said that all that he has offered is "the scientific explanation of the reliability of learning", with the implication that "such a scientific explanation must be compatible with any metaphysical position."[64] This raises what is essentially the same question as the one asked at the beginning of the second Locke Lecture - what is the bearing of quasi-scientific explanatory hypotheses pertaining to the "success" of certain human practices on metaphysical debates about the nature of truth? And Putnam offers what is essentially the same answer as the one which is given in the Locke Lectures. Given the availability of a quasi-intuitionist reading of the logical connectives, it seems that

one could take over the entire causal explanation of the
contribution of linguistic behaviour to the success of
total behaviour, without really being committed to the
existence of most of the entities science talks about,
or the existence of a correspondence between words and
such entities. Causal realism within science would appear
to be compatible with an idealist reinterpretation of
total science.   65

The problem, however, is that the idealist cannot preserve certain

"modal" statements which are themselves part of science. For example,

it is part of our "causal explanation" of learning that we may make mis-

takes, even in cases of perceptual learning. Our theory thus implies

statements of the form, "'S* might be warrantedly assertible even though

not-S". But such statements are not preserved, Putnam maintains, on a

quasi-intuitionist account which takes truth to be a form of warranted

assertibility. The general structure of Putnam's argument for this con-

tention - he terms it the "idealistic fallacy" argument - is as follows:

> For any predicate P the idealist may want to substitute
> for "true" one can find a statement S such that
> (8) S might have property P and still not be true
> follows from our causal theory of learning. And this is
> so simply because the "slop" between being warrantedly
> assertible, no matter how construed, and being true
> (assuming only Criterion T) is itself "built in" to
> our causal theory.   66

The argument, in other words, is that if one takes science seriously,

"at face value", or if one wishes to preserve its overall integrity,

one is committed to a realist view of truth because such a view is im-

plied by science so taken.

Endnotes - Chapter One

1. Putnam, 1975a, p. vii.

2. Putnam, 1975b, p. viii.

3. Putnam, 1985a.

4. Putnam, 1975a, p. vii.

5. Putnam, 1971, p. 347.

6. Ibid., p. 356.

7. Putnam, 1960.

8. Putnam, 1978, p. 4.

9. Putnam, 1975b, p. vii.

10. Ibid., pp. viii-ix.

11. Ibid., p. ix.

12. Ibid., p. x.

13. Kripke, 1972.

14. Putnam, 1983d, p. 71.

15. Putnam, 1975c.

16. Putnam, 1975a, p. xi.

17. Putnam, 1971, p. 350.

18. Putnam, 1962.

19. Putnam, 1975a, p. xi.

20. Ibid., p. xiii.

21. Putnam, 1978, p. 1.

22. Tarski, 1951.

23. Putnam, 1978, p. 3.

24. Ibid., p. 4.

25. Putnam, 1975e.

26. Putnam, 1975a, pp. vii-viii.

27. Putnam, 1975b, p. xiv.

28. Putnam, 1975a, p. xiii.

29. Putnam, 1975d, p. 93.

30. Ibid., p. 90.

31. Field, 1972.

32. Devitt, 1983a.

33. Putnam, 1978, p. 2.

34. Field, 1972, p. 362.

35. Putnam, 1978, pp. 4-5.

36. Putnam, 1976a, p. 16.

37. Ibid., p. 17.

38. Ibid., pp. 18-19.

39. Ibid., p. 19.

40. Ibid., pp. 20-21.

41. Ibid., p. 29.

42. Loc. cit.

43. Putnam, 1976a, pp. 34-35.

44. Ibid., pp. 35-36.

45. Ibid., p. 25.

46. Ibid., p. 37.

47. Ibid., pp. 50-51; p. 54.

48. Putnam, 1975a, p. xii.

49. See Putnam, 1976a, pp. 40-41, for this argument.

50. Putnam, 1976a, pp. 32-33.

51. Ibid., p. 41.

52. Ibid., p. 56.

53. Ibid., p. 57; p. 59.

54. Ibid., pp. 60-65.

55. Ibid., p. 57.

56. Ibid., pp. 57-58.

57. Putnam, 1978, p. viii.

58. Dummett, 1976c.

59. Putnam, 1976b, p. 110.

60. Ibid., pp. 98-99.

61. Ibid., pp. 110-111.

62. Ibid., pp. 100-101.

63. Ibid., p. 105.

64. Ibid., p. 107.

65. Ibid., p. 108.

66. Ibid., p. 108-109.

Chapter Two   "Metaphysical" and "Internal" Realism

II-a   The Jerusalem Exchange

In chapter one, I attempted to provide a reasonably comprehensive over-view of the relevant parts of Putnam's "pre-critical" writings. One reason for dwelling in some detail upon "sophisticated realism" is that an understanding of the latter position will greatly facilitate the comprehension of Putnam's more recent writings, especially the paper "Realism and Reason". A second reason, running almost contrary to the first, is that the critical discussion in later chapters will have little to say concerning the "pre-critical" works, but will focus, rather, upon arguments against one of two kinds of "realism" which Putnam distinguishes in responding to Dummett's commentary on "Reference and Understanding". Thus fullness of presentation in chapter one ran little risk of redundancy or repetition. The same does not apply to the material under review in this and the succeeding chapter, however, and I shall accordingly present some of the arguments in barest outline only, pending fuller treatment in the critical confines of later chapters.

Dummett's commentary on "Reference and Understanding"[1] contains three distinct critical responses to Putnam's defence of realism, the last of which proves to be the most significant. Firstly, Dummett contends that the "Idealistic Fallacy" argument has force only against an account of truth - "a theory about what, in general, makes a sentence true" - which identifies truth with a property that may be lost; it therefore has no force against "one who says that a mathematical statement, if true at all, is true only in virtue of the existence of a proof of it,"[2] for example. Secondly, he maintains that Putnam has not

shown that we require a "realist" account of truth to explain the contribution of language to the successful prosecution of our cognitive and non-cognitive goals. If we were able to construct "a workable non-realist semantics for our language,...this might serve equally well to explain our success, since any workable semantics must yield the result that a great deal of what we take to be true is in fact true."[3] Thirdly, Dummett questions whether "a truth-conditional semantics escape(s) the objections which face it as a theory of understanding by being transposed into an explanatory key". For, if we are to employ the "realist" notion of truth in explaining the contribution of language to the success of behaviour, we "must be able to acquire the notion somehow." But, so Dummett maintains, the very difficulties which plague the attempt to give a truth-conditional account of speakers' mastery of their language also plague the claim that we can avail ourselves of a realist notion of truth.

Putnam himself, we may recall, argues that the latter considerations do not pose a problem as long as the realist eschews a truth-conditional semantics. Dummett characterises Putnam's argument as follows:

> Hilary's view is that a Tarskian truth-definition will suffice to introduce [the realist notion of truth]. The objection that such a truth-definition, to be understood, requires us to already understand the language, is powerless in this context: according to Hilary's account, we do already understand the language, namely, by a mastery of its use. To yield the required realist semantics, the metalanguage in which the definition is given must have a classical logic; but there will be no difficulty in this, on Hilary's account, since we do already understand the classical logical constants simply by having been trained to perform deductions according to the classical laws. [4]

Dummett offers two criticisms of this argument. In the first place, he disputes the adequacy of Putnam's account of linguistic understanding, and thus the claim that a mastery of "use", in the specified sense, constitutes an understanding that is sufficient to ground the introduction of a "realist" notion of truth. A speaker who possesses only "a mastery of a number of interlocking practices connected by no unifying principle in terms of which it can be stated in what a grasp of the content of the sentence consists, and by reference to which the various linguistic practices...can be accounted for and justified...has, as it were, no grasp of what a sentence says."[5] Dummett rejects Putnam's account of linguistic understanding because of its holistic nature, and argues instead for a theory of understanding that identifies mastery of a language with a grasp of conditions of verification (or falsification) for individual sentences.[6]

Secondly, he claims that the combination of a Tarskian truth-definition and classical logic will not yield the required "realist" semantics within which truth is taken to satisfy bivalence. Although it might seem that such resources would allow us to derive the principle of bivalence in the meta-language, this is deceptive: for such a derivation would prove too much in that it would apply even to vague sentences, our use of which clearly indicates that we do not take them to be subject to bivalence. In fact, so Dummett argues, acceptance of a Tarskian truth-definition and of classical logic is quite compatible with a non-realist account of truth. In this case, the question arises how our use of language could ever manifest our grasp of a "realist" notion of truth. He concludes that "no notion of truth satisfying bivalence could have the required connection with use."[7] (See chapter six below for a

fuller discussion of Dummett's "manifestation argument" to which a tacit appeal is made in the foregoing). And, if we cannot acquire the "real-'ist" notion of truth, then we obviously cannot employ such a notion in an explanation of linguistic activity.

In his "Reply to Dummett's Comments"[8], Putnam addresses the first and the third of Dummett's criticisms. In response to the first criticism, he argues that, apart from mathematical sentences, verification or warranted assertibility is "a property that may be lost". In response to the third objection, he begins by charging that his position has been misrepresented: "My position is not that the laws of logic determine our concept of truth, but that our entire theory, including our theory of the relation of language to the speaker's environment, determines our concept of truth."[9] But he then concedes that this does not entirely deflect Dummett's criticism:

> The entire language may have - indeed does have - many
> interpretations, both in the model-theoretic sense of
> "interpretations", and in the more radical sense in which
> the logical constants themselves are allowed to be re-
> interpreted. And in what intelligible sense of "intended"
> is just one of these interpretations the "intended" one?[10]

This extremely telegraphic argument prefigures the "model-theoretic" argument of "Realism and Reason" and "Models and Reality". Putnam's point seems to be that "our entire theory" is no better able to ground the ascription to us of a "realist" notion of truth than is the combination of classical logic and a Tarskian truth-definition. If understanding a language consists in the ability to use that language, then our understanding of our entire language, manifested in the totality of language use, will not serve to manifest that we have a "realist" notion

of truth; nor, as Putnam will later argue, does such use serve to pick
out a unique relation of reference between language and 'the world'.

Putnam's response to these difficulties is to draw a distinction
between two kinds of realism, which he christens "metaphysical" and
"internal" realism. Internal realism is characterised as

> a first-order theory about the relation of a language
> (actually, of the speakers of a language) to the speak-
> er's environment. From within such a story, the notion of
> a "correspondence" between words and sets of things is
> as legitimate and meaningful as the notion of a chair
> or a pain. What is well-defined and what is not, how
> two-valued truth explains the success of language-using,
> are all parts of the story (on all fours with the beha-
> viour of electricity or the properties of heat).   11

This is to be contrasted with metaphysical realism, which is

> a picture (or a "model", in the sense in which collid-
> ing billiard balls are a "model" for a gas) of the
> relation of any correct theory to The World. The pic-
> ture is that each term in a correct theory is a label
> for a determinate piece (or kind of piece) of The World. 12

Putnam's contention is that, while considerations raised by Dummett and
Nelson Goodman may indeed demonstrate the incoherence of metaphysical
realism, this in no way invalidates internal realism. The internal
realist has an easy answer to Dummett, Putnam claims, namely, that
"what is 'vague' and what is not, what is an 'intended' interpretation
and what is not, are questions with no absolute sense. They can only be
answered from the standpoint of one or another meta-theory."[13] How this
is supposed to work is something that receives further clarification in
the paper, "Realism and Reason", to which we should now turn our atten-
tion.

II-b   "Realism and Reason"

"Realism and Reason", written in the fall of 1976 and presented at the December meetings of the A.P.A., is the paper in which Putnam first attempts to spell out in detail the distinction between the two kinds of realism noted in the previous section, and to offer arguments for retaining internal realism while rejecting metaphysical realism. In the Preface and Introduction to MMS, he characterises "Realism and Reason" as a "recent turn" in his thinking, and provides the following already cited remarks concerning the general import of the paper:

> Broadly speaking, the thrust is that a "verificationist" view of truth and a correspondence theory of truth (in the sense of empirical realism) are not incompatible. Inasmuch as Kant had a view which included a correspondence view of truth within the empirical realm (on my reading anyway) and a stress on the mind-dependence of all truth, there is (I like to think) something Kantian in the view with which I end up. But it is a demythologised Kantianism, without "things in themselves" and "transcendental egos". [14]

Despite these helpful "reader's guides", commentators have found "Realism and Reason" a far from transparent document.[15] The exegetical difficulties begin on the very first page, where Putnam draws once again the distinction between the two kinds of realism. Internal realism, we are told, is an "empirical theory" which explains such facts as that "scientific theories tend to 'converge'", and that "language-using contributes to getting our goals, achieving satisfaction, or what have you." The internal realist explains such facts by hypothesising that "speakers mirror the world - i.e. their environment - in the sense of constructing a symbolic representation of that environment...A 'correspondence' between words and sets of things...can be viewed as part

of an _explanatory model_ of the speakers' collective behaviour." Internal realism, consisting in "acceptance of this sort of scientific picture of the relation of speakers to their environment, and of the role of language," is, so Putnam claims, the view spelled out and argued for in the _Locke Lectures_ and in "Reference and Understanding".[16]

It would seem, from these remarks, that internal realism is simply the position argued for in the "pre-critical" writings. But, in that case, what is "metaphysical" realism? And where is the "new turn" in Putnam's thinking? And how are we to account for certain observations later in "Realism and Reason" to the general effect that the "pre-critical" writings were an expression of _metaphysical_ realism? For example, Putnam states that his attempt, in "Reference and Understanding", to carve out an explanatory role for the realist notions of truth and reference while embracing a form of "verificationist semantics" gave Dummett "all he needs to demolish metaphysical realism - a picture I was wedded to!"[17] Or, again, he talks of his earlier attempts to defend _metaphysical_ realism by appeal to the doctrine of "invariants" and "equivalent descriptions".[18] Such considerations have prompted Michael Devitt to question the intelligibility of the supposed "change" in Putnam's position. Devitt distinguishes two theses that he takes to be "realist" in nature:

> "Physical Realism" (PR): Common sense, and scientific, physical entities objectively exist independently of the mental.
> The "Correspondence Notion of Truth" (CNT): Physical statements are true or false in virtue of: (i) their structure; (ii) the referential relations between their parts and reality; and (iii) the objective nature of that reality.

Devitt then expresses his exegetical puzzlement as follows:

> As I understand it, ["Reference and Understanding"]
> argues that the combination of world view and notion of
> truth that best explains the success of linguistic be-
> haviour, preserves certain deductive rules, and explains
> the reliability of our learning, is a realist one, a
> combination like PR and CNT. If this empirical theory
> were "empirical realism", all well and good. But then
> this theory is "metaphysical realism", which he now
> rejects. I am at a loss to know how we can subtract
> "metaphysical realism" from this and still be left with
> anything, let alone anything that is appropriately called
> a "correspondence" notion of truth and "realism" (even
> if only of the "internal" variety.   19

To relieve Devitt's puzzlement, we need to answer at least the
following questions: (a) What is the relationship between "metaphysical"
and "internal" forms of realism, and (b) how, in terms of these two
forms of realism, are we to characterise Putnam's "pre-critical" position
so as to make intelligible the purported change in his views? At least
a partial answer to these questions may be forthcoming if we attend more
closely to the characterisation of "metaphysical" realism. As in his
reply to Dummett, Putnam describes the latter as "less an empirical
theory than a model": "It is, or purports to be, a model of the relation
of any correct theory to all or part of THE WORLD." He then sketches
the model in question:

> In its primitive form, there is a relation between each
> term in the language and a piece of THE WORLD (or a kind
> of piece if the term is a general term).
> This relation, the relation of reference, is given by
> the truth-conditional semantics for the language, in the
> canonical versions of the theory - i.e. understanding a
> term, say T$_j$, consists in knowing what piece of THE WORLD
> it refers to... Minimally, however, there has to be a
> determinate relation of reference between terms in L and
> pieces (or sets of pieces) of THE WORLD, on the meta-
> physical realist model, whether understanding L is taken
> to consist in "knowing" that relation or not. What makes

> this picture different from <u>internal</u> realism (which
> employs a similar picture <u>within</u> a theory) is that
> (1) the picture is supposed to apply to <u>all</u> correct
> theories at once (so that it can only be stated with
> "typical ambiguity" - i.e. it transcends complete for-
> malisation in any one theory); and (2) THE WORLD is
> supposed to be <u>independent</u> of any particular representation
> we have of it - indeed, it is held that we might be <u>un-
> able</u> to represent THE WORLD correctly at all.   20

Metaphysical realism, then, purports to provide a model of what it is

for any particular theory to be "correct". "Correctness" consists in a

reference-mediated correspondence to THE WORLD, and a theory which so

corresponds is <u>true</u>. In so far as this model is to serve for <u>all</u> theo-

ries, the notions employed in formulating the model - such notions as

reference, correspondence, truth, and THE WORLD - cannot receive their

interpretation from within any particular theory. In the "pre-critical"

writings, Putnam argues for "empirical realism" as a "correct" theory

in the <u>metaphysical realist sense</u>. So understood, "realism as an over-

arching empirical hypothesis" contends that speakers construct repre-

sentations of THE WORLD and that these representations <u>refer</u> to, and

are approximately <u>true of</u>, THE WORLD. To reject metaphysical realism

while retaining "internal" realism is therefore to uphold the theses

of "empirical realism" while refraining from interpreting those theses

in the foregoing manner

But how <u>are</u> we to interpret the theses of "empirical realism" if

we follow Putnam in taking the "internal turn"? Before addressing this

question, however, I shall briefly outline the principal argument that

Putnam offers against metaphysical realism in "Realism and Reason" and

in the paper "Models and Reality".

II-c   The "Model-Theoretic" Argument

What is usually termed the "model-theoretic" argument against
metaphysical realism receives two fairly lengthy expositions in Putnam's
writings. The first, relatively "popular", presentation of the argument
occurs in "Realism and Reason". A lengthier and more technical treatment
of the subject takes up the bulk of the later paper "Models and Reality".
In fact, much of the technical detail of the latter presentation is un-
necessary for our purposes, and, indeed, for Putnam's in so far as his
intention is to undermine metaphysical realism (see chapter five below
on this). It is noteworthy that, in his more recent writings, Putnam
has employed a much simpler version of the "model-theoretic" argument.[21]
I shall therefore rely mainly upon "Realism and Reason" in the present
context, referring to "Models and Reality" only in order to clarify
certain features of the argument. Such clarification is important if we
are to avoid some common misconceptions of the argument, as we shall
see in chapter five below.

Putnam contends that metaphysical realism is "incoherent". His
argument in support of this claim draws crucially upon two features of
the realist model. Firstly, there is the feature that is to be the
focus of Putnam's attack:

> The most important consequence of metaphysical realism is
> that truth is supposed to be radically non-epistemic - we
> might be "brains in a vat" and so the theory that is ideal
> from the point of view of operational utility, inner
> beauty and elegance, "plausibility", "simplicity",
> "conservatism", etc., might be false. "Verified" (in any
> operational sense) does not imply "true", on the meta-
> physical realist picture, even in the ideal limit.   22

Secondly, there is the feature implicit in the claim that even an
"Ideal Theory" might be false and already cited above, namely, that

there has to be a determinate reference relation between linguistic elements and entities in THE WORLD. Only if there is such a relation can states of THE WORLD confer determinate truth-values upon sentences containing such linguistic elements, and thus confer the determinate value "false" upon some sentence belonging to an "Ideal Theory".

The argument. proceeds by way of (i) a specification of what we are to understand by the term "Ideal Theory", (ii) a demonstration that there is an interpretation of the language of the Ideal Theory such that all the sentences that the latter comprises come out true, and (iii) the contention that we cannot exclude this "interpretation" as in some way "unintended" or otherwise illegitimate. I shall reproduce the salient passages from the text for future reference:

> Let T_I be an ideal theory, by our lights. Lifting restrictions on our all-too-finite powers, we can imagine T_I to have every property except objective truth - which is left open - that we like. E.g. T_I can be imagined complete, consistent, to predict correctly all observation sentences (as far as we can tell), to meet whatever "operational constraints" there are (if these are "fuzzy", let T_I seem clearly to meet them), to be "beautiful", "simple", "plausible", etc. The supposition under consideration is that T_I might be all this and still be (in reality) false. I assume THE WORLD has (or can be broken into) infinitely many pieces. I also assume T_I says there are infinitely many things (so in this respect T_I is "objectively right" about THE WORLD). Now T_I is consistent (by hypothesis) and has (only) infinite models. So by the completeness theorem (in its model theoretic form) T_I has a model of every infinite cardinality. Pick a model M of the same cardinality as THE WORLD. Map the individuals of M one-to-one into the pieces of THE WORLD, and use the mapping to define relations of M directly in THE WORLD. The result is a satisfaction relation SAT - a "correspondence" between the terms of L and sets of pieces of THE WORLD - such that the theory T_I comes out true - true of THE WORLD - provided we just interpret "true" as TRUE(SAT). So what becomes of the claim that even the ideal theory T_I might really be false? .23

Putnam considers two possible responses to this question. Firstly, it might be argued that "SAT is not the _intended_ correspondence between L and THE WORLD". The problem here, however, is to point to some constraints on reference that can serve both to select a unique "intended" interpretation of L and to justify the rejection of SAT as "unintended". For SAT satisfies all "operational constraints", and thus SAT will certainly _seem_ to be true. And SAT also "certainly meets all _theoretical_ constraints on reference - it makes the _ideal_ theory, T_I, come out true." Secondly, an appeal might be made to a "causal theory" of reference. Putnam rejects this suggestion on the grounds that "how 'causes' can uniquely refer is as much of a puzzle as how 'cat' can, on the metaphysical realist picture."[24]

Putnam's presentation in "Models and Reality" may help to illuminate certain features of the argument. The initial focus in this later paper is the Skolem Paradox in model theory and the problem of singling out a unique "intended model" for set theory. Neither operational nor theoretical constraints, it is argued, can capture our intuitive notion of "set", and, for a naturalistic philosopher who identifies understanding a language with the ability to _use_ that language, it is difficult to see what further constraints are available. Putnam then argues for a number of extensions of this result. In each case, the thrust of the argument is that, given a set of operational constraints formulated in terms of _one_ vocabulary - relatively speaking, the "observational" vocabulary - we cannot single out a unique, "interpretation-independent" reference for the terms in another vocabulary - relatively speaking, the "theoretical" vocabulary - by means of such operational constraints in conjunction with a set of theoretical constraints. In the first

instance, the "theoretical" vocabulary consists of terms purporting to refer to theoretical entities in science, while the "observational" vocabulary consists of terms that refer to things, events, and properties "observable with the human sensorium". In the second instance, the "observational" vocabulary is a sense-datum language, and the "theoretical" vocabulary includes such terms as 'cat' and 'dog'. Finally, we may take the "observational" vocabulary to comprise only terms that refer to present sense-data, and include all terms that refer to past and future sense-data in the "theoretical" vocabulary. In each of these cases, we can run what is essentially the argument of "Realism and Reason" to establish two conclusions: (i) an "ideal theory" formulated in terms of the "theoretical" vocabulary cannot intelligibly be supposed to be false, and (ii) terms in the "theoretical" vocabulary do not get assigned a unique and determinate reference, and must be regarded as "formal constructs variously interpreted in various models."[25]

We may note one further point concerning the "model-theoretic" argument in its various incarnations, namely, that it is an attempt to undermine metaphysical realism from within, to demonstrate that the realist model is "internally incoherent". The discussion proceeds in terms of the very realist notions that it seeks to discredit - the notions of truth and reference construed as modes of "correspondence" to THE WORLD. The contention is that there are not adequate resources, within the metaphysical realist model, to make sense of the distinctive realist thesis that even an "ideal theory" might be false.

II-d Internal Realism

It was suggested above that the crucial difference between the
"sophisticated realism" ofthe "pre-critical" writings and the "internal"
realism espoused in the opening paragraphs of "Realism and Reason" is
that, while both forms of realism are characterised as "over-arching
empirical hypotheses" pertaining to the abilities of speakers of a
language to construct a "symbolic representation" of their environment,
the central terms employed in the formulation of the realist hypothesis
receive a "metaphysical realist" interpretation in the former case and
an interpretation from "within our theory" in the latter case. It was
also suggested that the notions of truth and reference that emerge on
such an "internal" interpretation stand in need of further clarificat-
ion. In seeking such clarification, we enter one of the more impenet-
rable thickets of Putnamian exegesis. For the very thesis of "internal
realism" seems to undergo subtle transformation within the pages of
"Realism and Reason". We begin with the idea that internal realism is
a "scientific picture of the relation of speakers to their environment,
and of the role of language," a picture that we should endorse because
of its capacity to explain certain features of scientific methodology
and the general success of linguistically-mediated endeavours. In
"Reference and Understanding", as we have seen, it is argued that this
"empirical theory", with its implicit "realist" notion of truth, is
compatible with the acceptance of a non-realist semantics. But, as we
shall see, when Putnam argues that internal realism is not vulnerable
to the "model-theoretic" argument, the crucial feature of the former
upon which the argument rests is its incorporation of a non-realist
semantics. Indeed, in both "Realism and Reason" and "Models and Reality",

we seem to be presented with a choice between a metaphysical realism that incorporates a truth-conditional semantics, and an internal realism that incorporates a verificationist semantics. But what has become of the idea, defended in "Reference and Understanding", that one can combine a verificationist-type semantics with a form of "sophisticated realism"? In a puzzling passage, Putnam rehearses the argument of "Reference and Understanding" and then comments that "now it looks as if in conceding that some sort of verificationist semantics must be given as our account of understanding (or "linguistic competence", in Chomsky's sense), I have given Dummett all he needs to demolish metaphysical realism - a picture I was wedded to!"[26] But the immediately preceding discussion is to the effect that the problem posed by the "model-theoretic" argument does not arise if one follows Dummett in espousing a non-realist semantics. Since Putnam already grants the case for a non-realist semantics in "Reference and Understanding", how do these considerations enable Dummett to "demolish metaphysical realism"?

I have already ventured some thoughts on these matters in section II-a above. Rather than rehearse those observations, therefore, I shall turn to the text of "Realism and Reason" and "Models and Reality" for further illumination. One issue, canvassed at length in both papers, is the bearing of the "model-theoretic" argument on one who espouses a non-realist semantics. Putnam, citing Dummett, characterises a "non-realist semantics" as a semantic theory "which holds that a language is completely understood when a verification procedure is suitably mastered and not when truth-conditions (in the classical sense) are learned."[27] The critical feature in Dummett's proposed semantics, serving to distin-

guish.it from traditional phenomenalist positions, is that there is "no 'basis' of hard facts (e.g. sense data) with respect to which one ultimately uses the truth conditional semantics, classical logic, and the realist notions of truth and falsity:"

> If sense data are treated as "hard data" - if the verificationist semantics is given in a meta-language for which itself we give the truth-conditional account of understanding - then we can repeat the whole argument against the intelligibility of metaphysical realism (as an argument against the intelligibility of the meta-language) - just think of the past sense data (or the future ones) as the "external" part of THE WORLD...This is why Dummett's move depends upon using the verificationist semantics all the way up (or down) - in the meta-language, the meta-meta-language, etc.   28

This is essentially a variant on the "Models and Reality" version of the "model-theoretic" argument. Suppose it is claimed that one understands a sense-datum statement by grasping its (realist) truth-conditions. It follows, clearly, that one can understand only those sense-datum statements that possess determinate realist truth-conditions. But we may take discourse about present sense-data as our "observational" vocabulary, in terms of which we may formulate "operational constraints" governing the interpretation of discourse about past and future sense-data, taken as discourse involving our "theoretical" vocabulary. Our use of the language will then, by the "model-theoretic" argument, fail to pick out a unique interpretation of the "theoretical" vocabulary, and, as a consequence, past and future sense-datum statements will lack determinate realist truth-conditions, and therefore, ex hypothesi, be incapable of being understood. And, if such statements belong to the meta-language in terms of which we formulate our understanding of the object language, this unintelligibility will infect the language in

general.

.We might enquire, however, what it means to use the verificationist
semantics "all the way up (or down)". Putnam offers the following eluc-
idatory remarks:

> Understanding a sentence, in this semantics, is knowing
> what constitutes a proof (verification) of it. And this
> is true even of the sentences that describe verifications.
> ...Whatever language I use, a primitive sentence - say,
> "I see a cow" - will be assertible if and only if verif-
> ied. And we say it is verified by saying the sentence
> itself, "I see a cow"..."I see a cow" is "self-warranting"
> in this kind of epistemology - not in the sense of being
> incorrigible, not even necessarily in the sense of being
> fully determinate (i.e. obeying strong bivalence - being
> determinately true or false)...The important point is
> that the realist concepts of truth and falsity are
> not used in this semantics at all.   29

The general import of these remarks has proved elusive for at least
one commentator:

> I find this obscure. Notice, the passage begins by ex-
> plaining assertibility in terms of verification. But then
> verification for basic or primitive sentences is explain-
> ed in terms of what? Apparently in terms of assertion!
> But clearly the same "explanation" is available to the
> realist. The primitive sentence, "I see a cow", will be
> assertible iff true. And we say it is true by saying the
> sentence itself. The important point is that the anti-
> realist concepts of verification and falsification are
> not used in this semantics at all. So what has been
> shown? I think only that the assertibility condition
> for a sentence may be given in terms of the underlying
> meaning theory: Verification and falsification, in the
> one case, and truth and falsity in the other. But the
> verification condition for a sentence is just as much
> extra-linguistic for the anti-realist as the truth-
> condition is extra-linguistic for the realist. How we
> make the assertion that s is verified - or, for that
> matter, how we make the assertion that s is true -
> sheds no light on what its verification or truth con-
> sists in - unless, of course, one adopts the view that
> the verification of s consists in its assertion by us.
> But this is simply not a notion of verification at all;
> even if verification involves agreement, it is certainly

Demopoulos' comments raise a number of issues. An obvious response
to the argument contained in the first part of the quoted passage is
that the availability of a parallel "realist" account of the assert-
ibility of basic sentences fails to undermine Putham's argument. For
the significance of the claim that the non-realist semantics does not
require the realist notions of truth and falsity, even for basic sen-
tences, is that, as we have seen, the "model-theoretic" argument seems
to demonstrate that any semantic theory that _does_ require these realist
notions renders "unintelligible" our capacity to understand language.
But Demopoulos also raises a more troubling question, namely, how are
we to make sense of the claim that understanding even basic sentences
is a matter of knowing verification conditions? He reads Putnam as
saying that "verification", for such sentences, is to be explained in
terms of how we _assert_ them to be verified; and he has difficulty,
understandably enough, making sense of such a proposal. But this read-
ing is surely mistaken. Putnam's point is _not_ that, for a given basic
sentence _s_, _s_'s being verified somehow _consists_ in our _asserting_ _s_,
but, rather, that the verification conditions for a basic sentence can
be given only by reiterating the sentence itself. The "self-warranting"
nature of such sentences comes out in the fact that their verification
conditions admit of only "trivial" formulations, of the form "'_s_' is
verified iff s"..

On the proposed reading, Putnam is innocent of Demopoulos' charge
because he doesn't even _address_ the question, "In what does the verif-
ication of basic sentences consist?", in the passage under consideration.

But it seems fairly clear how he would propose to answer that question. The verification condition for a basic sentence is presumably the occurrence of the sort of experience that warrants the assertion of the sentence, given the practice of the linguistic community, and to assert the sentence is itself to assert that the appropriate verification condition obtains. A certain degree of caution is necessary here, however, given Putnam's talk of "sentences that describe verifications". For, if we recall Demopoulos' observation that verification conditions and truth conditions are equally extra-linguistic, we may wonder whether our understanding of "sentences that describe verifications" presupposes a unique "reference" relation between terms occurring in those sentences and features of the extra-linguistic verifications that they describe. We should surely resist such a conclusion, if possible, given Putnam's repeated insistence that the non-realist semantics completely eschews realist notions of truth and reference. The phrase, "sentences that describe verifications", is unfortunate in its suggestion that we require a referential semantics for such sentences. Putnam's purposes might be better served by a term such as "verification-description", understood as an unbreakable predicate ranging over sentences (c.f. "unicorn-picture"). To understand a verification-description is to know under what circumstances the assertion of a particular sentence is warranted, and this presupposes a grasp of realist truth conditions neither more nor less than does the understanding of any other sentence within the framework of a non-realist semantics.

As a further corrective to the temptation to regard the understanding of basic sentences as somehow requiring a realist correspondence between language and features of an extra-linguistic reality, we

might note that there is no requirement, on a non-realist semantics, that speakers who share a common understanding of the basic sentences of a language also share verifying "experiences", phenomenologically conceived. Suppose, for example, that the experience that serves as a verifying condition for the basic sentence, "I see something green", for one speaker A is qualitatively identical to the experience that serves as a verifying condition for the basic sentence, "I see something purple", for another speaker B and vice versa. Despite this phenomenological disparity, A and B can still agree with one another and with other members of the linguistic community as to when each sentence is properly assertible. And, to the extent that they do so agree, each will understand the relevant fragment of the language according to the non-realist semantics. But our account of understanding neither employs, nor has any place for, the idea that there must be a unique relation of reference between the use of a given term by A and B and some feature of an extra-linguistic reality.

I have dwelt on these issues at some length because I think that they lend themselves to misunderstanding. Now, however, we may return to the question, how does the "model-theoretic" argument bear upon the proponent of a non-realist semantics? In "Models and Reality", Putnam contends that such an individual evades the clutches of the "model-theoretic" argument because such a semantics provides "a standpoint which links use and reference." The failing of a truth-conditional semantics is precisely its inability to provide such a standpoint. For on such a semantics the interpretation of a language is given by specifying a set of "intended" models, and a speaker's understanding of the language - his knowledge of the truth conditions for sentences in the language -

consists in his associating elements in the language with their inter-
pretations in the set of models. The _problem_ for the realist, according
to Putnam, is that use of the language, as captured by operational and
theoretical constraints, fails to pick out the set of "intended" models
that the realist requires. The non-realist semantics, on the other hand,
makes no essential use of the notion of a model, or, _à fortiori_, of the
notion of an "intended" model: a sentence is understood when the appro-
priate verification procedure is mastered. In the case of a non-realist
semantics, therefore, "the 'gap' between words and world, between our
use of language and its 'objects' never appears."[31] The problem with
the realist semantics is that a complete account of use leaves the _inter-
pretation_ of the language, and thus its understanding, indeterminate.
But, on the non-realist semantics, the interpretation of a language is
given _through_ its use, and understanding _consists_ in a mastery of use.

It might be thought, however, that one can still _introduce_ the
"realist" notions of truth, reference, and model into the language,
and that, in so far as one does so, one will face the same difficulties
seen to attend the realist semantics. Of course, the strategy of adop-
ting a non-realist semantics as an account of linguistic understanding
and then maintaining that we can _acquire_ the realist notions of truth
and reference and employ these notions to formulate a kind of realism --
"realism as an empirical hypothesis" - is precisely the strategy of
"Reference and Understanding". Given that "Realism and Reason" is at
best suggestive as to the infirmities of the "Reference and Under-
standing" strategy, the brief, if unacknowledged, discussion of the
latter in "Models and Reality" is of special interest. Putnam begins by
noting that the proponent of a non-realist semantics can indeed avail

himself of the notions of truth, reference, and model, even though these

notions are not used in his semantics:

> He has to foreswear reference to models in his account
> of understanding; but, once he has succeeded in under-
> standing a rich enough language to serve as a meta-
> language for some theory T (which may itself be simply
> a sub-language of the meta-language, in the familiar
> way) he can define "true in T" á la Tarski, he can talk
> about "models" for T, etc. He can even define "reference"
> or "satisfaction" exactly as Tarski did.   32

Putnam then considers whether the proponent of a non-realist seman-

tics who pursues such a course of action will thereby be forced to con-

front the "Skolem paradox" as exemplified in the "model-theoretic" argu-

ment. He addresses the particular case of mathematical intuitionism,

rather than a more general non-realist semantics, but we may extract a

moral for the general case from his treatment of the particular:

> "Objects" in constructive mathematics are given through
> descriptions. Those descriptions do not have to be mys-
> teriously attached to those objects by some non-natural
> process (or by metaphysical glue). Rather, the possibil-
> ity of proving that a certain construction (the "sense",
> so to speak, of the description of the model) has certain
> constructive properties is what is asserted and all that
> is asserted by saying that the model "exists". In short.
> reference is given through sense, and sense is given
> through verification procedures and not through truth
> conditions. The "gap" between our theory and the "objects"
> simply disappears - or, rather, it never appears in the
> first place.   33

Putnam's point is that meta-mathematical talk of "models" is itself

understood constructively within intuitionistic mathematics. This is

to reaffirm, for the particular case, the non-realist contention that

the verificationist semantics has to be applied "all the way up (or

down)." Extrapolating to the general case, the proponent of a non-

realist semantics can indeed talk in the metalanguage about the truth conditions of sentences and the reference of terms, and he can even characterise linguistic understanding in terms of the knowledge of truth conditions. As Putnam says, "the 'non-realist' semantics is not inconsistent with realist semantics; it is simply prior to it, in the sense that it is the 'non-realist' semantics that must be internalised if the language is to be understood."[34] But all such talk in the metalanguage will itself be understood in terms of the appropriate verification conditions. We may introduce 'true' a la Tarski, but the metalanguage in terms of which we formulate our 'truth-definitions' receives its interpretation via the non-realist semantics, and the notion of truth thereby introduced will not be the metaphysical realist notion.

The failure of the "Reference and Understanding" strategy may therefore be explained as follows: (a) in order for the non-realist semantics to be a coherent option, it must be employed "all the way up (or down)", but (b) if it is so employed, we cannot introduce the requisite realist notions of truth and reference into our language: the only notions of truth and reference that we can understand will fail to satisfy the realist's requirements. For similar reasons, a non-realist semantics can permit the introduction of notions of truth and reference in the metalanguage without falling prey to the "model-theoretic" argument, for the latter can bite only where we have genuinely realist notions of truth and reference. The points are well illustrated in the closing paragraphs of "Models and Reality", where Putnam considers how the non-realist semantics fares in the face of the claim that there might be a model of our language that preserved all operational and theoretical constraints yet permuted the extensions of the terms 'cat' and

'dog', and that this model would be just as "intended" as the unpermuted one:

> Such a model would be unintended <u>because we do not in-</u>
> <u>tend the word 'cat' to refer to dogs.</u> From the meta-
> physical realist standpoint, this answer does not work;
> it just pushes the question back to the metalanguage.
> The axiom of the metalanguage, "'cat' refers to cats",
> cannot rule out such an unintended interpretation of
> the object language, unless the metalanguage itself
> already has had <u>its</u> intended interpretation singled out;
> but we are in the same predicament with respect to the meta-
> language that we are in with respect to the object language,
> from that standpoint, so all is in vain. However, from
> the viewpoint of "non-realist" semantics, the meta-
> language is completely understood, and so is the object
> language. So we can <u>say and understand</u> "'cat' refers
> to cats". Even though the model referred to satisfies
> the theory, etc., it is "unintended" <u>from the descrip-</u>
> <u>tion through which it is given</u> (as in the intuitionist
> case). Models...are constructions within our theory
> itself, and they have names from birth.   35

In "Realism and Reason", we encounter an argument similar to the foregoing, but offered in response to the suggestion that the "model-theoretic" argument refutes <u>internal</u> realism. As mentioned at the begin-ning of this section, the "internal realism" defended in the later pages of "Realism and Reason" seems to be identified more with the non-realist semantics of "Reference and Understanding" than with the so-called "empirical realism" of that paper. The internal realist is pres-ented with the following challenge: "How do you know that 'cow' refers to cows in the sense of referring to one determinate set of things, as opposed to referring to a determinate set of things <u>in each admissible</u> <u>interpretation</u>?", given that "there are other interpretations of your whole language...which would render true an <u>ideal</u> theory (in your lan-guage)."[36] Part of the answer to this challenge, Putnam contends, is that "the question whether the theory has a <u>unique</u> intended interpret-

ation has no absolute sense":

> Viewed from within Story 1 (or a metalanguage which
> contains the object language of Story 1), ['cow'] has
> a "unique intended interpretation". Viewed from within
> Story 2 (or a metalanguage which contains the object
> language of Story 2), the term ['cow'] as used in Story 1
> has a plurality of admissible interpretations. The
> critic's "how do you know?" question assumes a theory-
> independent fact of the matter as to what a term in a
> given theory corresponds to - i.e. assumes the picture
> of metaphysical realism.   37

The distinction Putnam is drawing here is between (a) cases where the
term to be interpreted belongs to the language in which the interpret-
ation is to be given, and (b) cases where this is not so. (This is
essentially Quine's distinction between translation situations where
we "acquiesce" in our "home" language, and translation situations where
we do not, or cannot, so proceed.[38]) In the second sort of case, there
is indeed a multiplicity of equally admissible ways of interpreting the
alien term into the "home" language. But, so it is claimed, the first
sort of case does admit of a unique "intended" interpretation on an in-
ternalist view of things. And since the "model-theoretic" argument
addresses our ability to understand our own language - i.e. the first
sort of case - it fails to undermine internal realism.

But why is the internal realist entitled to the luxury of a
"unique intended interpretation" in the first sort of case? Putnam's
answer to this question draws crucially upon the fact that internal
realism incorporates not just the "scientific picture" sketched at the
beginning of "Realism and Reason", but also the non-realist semantics
conceded to Dummett in "Reference and Understanding". He canvasses the
objection that a Tarski-like definition of 'reference' (or 'satisfac-

tion') in terms of a <u>list</u> of ordered pairs of the form $\langle$'cow', cows$\rangle$ – the sort of definition by means of which the internal realist can supposedly introduce the notion of reference – fails to <u>explicate</u> reference and, if anything, "<u>presupposes</u> that 'cow' refers (in some <u>other</u> sense of 'refers')."[39] This, as Putnam acknowledges, is Field's objection to Tarski that provides the focus for the <u>Locke Lectures</u>. But Putnam's response to this objection in "Realism and Reason" differs from his response in the latter context:

> The <u>use</u> of 'cow' does presuppose that 'cow' is <u>under-stood</u>. And if my account of <u>understanding</u> was a <u>truth-conditional</u> (or <u>reference-conditional</u>) account, then the objection would be good. But I gave a verificationist account of understanding (in terms of degree of confir-mation); thus <u>my use of the term 'cow' in the language</u> has already been explained, and I am free to use it – even to use it in explaining what 'cow' refers to. What I am saying is that, in a certain "contextual" sense, it is an <u>a priori</u> truth that the word 'cow' refers to a determinate class of things (or a more-or-less determinate class of things – I neglect ordinary vagueness). Adopting "cow talk" is adopting a "version", in Nelson Goodman's phrase, from within which it is a <u>priori</u> that the word 'cow' refers (and, indeed, that it refers to cows). [40]

We may paraphrase, and perhaps clarify, the foregoing in the following fashion. If we try to give a Tarski-like definition of 'ref-erence' in terms of a list – whether of ordered pairs or, less formally, of statements of the form, "'<u>T</u>' refers to T" – then each item in the list will contain two occurrences of a given term 'T'. In its first occurrence, 'T' is <u>mentioned</u>, and in its second occurrence it is <u>used</u> as part of the metalanguage. Our understanding of the definition of 'reference' obviously presupposes that we understand the metalanguage in which it is given, and thus that we understand the term 'T' as used

in that metalanguage. Now suppose, firstly, that one gives a realist

semantics as an account of what it is to understand the metalanguage -

an account whereby understanding the metalanguage is taken to consist

in interpreting it through one's grasp of the set of "intended" models

that fixes the reference of terms and the truth conditions of sentences.

In this case, it is obvious that the intelligibility of the proposed

definition of 'reference' presupposes a further sense of reference im-

plicated in the semantics for the metalanguage. Furthermore, given the

"model-theoretic" argument, we know that our entire use of the meta-

language fails to pick out unique extensions for its terms. Thus, while

we may claim to know that the sentence "'Cows' refers to cows" is true

in all admissible interpretations of the language, we cannot know that

"'cow' refers to cows in the sense of referring to one determinate set

of things, as opposed to referring to a determinate set of things in

each admissible interpretation;" the set of horse and the set of cows

in one admissible interpretation may be permuted in another equally

admissible interpretation. If, on the other hand, we adhere to a non-

realist semantics for the language as a whole, these difficulties do not

arise. One's understanding of both object- and meta-language consists

in one's mastery of the appropriate verification conditions, and one

can use the necessary terms in the metalanguage in one's definition of

'reference' without thereby presupposing a prior sense of reference.

And, in the absence of any such realist notion of reference, one cannot

mobilise the "model-theoretic" argument against the internal realist by

raising questions about what terms in the object language "really" refer

to in diverse "admissible interpretations". For, to the extent that we

adopt a particular "version" and raise the matter of the interpretation

of the terms of that version from <u>within</u> - i.e. to the extent that we are in the first sort of case distinguished above, where the object language is contained within the metalanguage - it is knowable <u>a priori</u> that, e.g. 'cows' refers to cows, because it follows immediately from the definition of 'refers' as a predicate that satisfies what Putnam terms Criterion S, the generalisation of Tarski's Criterion T to formulae containing free variables.[41] There is no further notion of reference available in terms of which we could give sense to the suggestion that 'cows' might really refer to horses, <u>unless</u> we shift to the second sort of case and regard 'cows' as an "alien" term.

Let me conclude by returning to our earlier inquiries as to the fate of the notions employed in the formulation of "empirical realism" - such notions as truth, reference, correspondence, and the world - when we take the "internal turn". A related matter is the validity of Devitt's charge that internal realism is really not a form of <u>realism</u> at all. I have suggested that the defence of internal realism in the later sections of "Realism and Reason" draws not upon the "scientific picture" sketched at the beginning of that paper, but, rather, upon the non-realist semantics first bruited by Putnam in "Reference and Understanding". But the <u>realist</u> pretensions of internal realism would seem to rest upon features of the "scientific picture" of "empirical realism" - the Boydian argument for convergence and the consequent "causal-explanatory" role accorded to the notions of truth and reference. To what extent does this "scientific picture" remain coherent when divorced from metaphysical realism and forced to cohabit with the non-realist semantics defended in the later sections of "Realism and Reason"?

In the first place, it is far from clear how we are to reconcile

the non-realist account of reference sketched in the preceding pages with the Boydian thesis that reference plays a "causal-explanatory" role in an explanation of the success of scientific practice. Boyd's contention, we may recall, is that the hypothesis that "terms in a mature science typically refer" is part of our explanation of scientific convergence. But suppose that we extrapolate from Putnam's observation that "adopting 'cow talk' is adopting a 'version'...from within which it is a priori that the word 'cow' refers (and that it refers to cows)." By analogy, it would seem, adopting "electron talk" is also adopting a "version" from within which it is a priori that 'electron' refers, and that it refers to electrons. Now contemporary science does adopt "electron talk", so it would seem that, talking from within contemporary science as good internalists should, we can know a priori that 'electron' refers (and refers to electrons). If we can know this a priori, however, it cannot also be an empirical hypothesis that helps to explain the success of the scientific enterprise. Or so it would seem.

Perhaps Putnam's answer to this sort of objection is implicit in his remark that

> we can revise "'cow' refers to cows" by scrapping the
> theory itself (or at least scrapping or challenging the
> notion of a cow) - and this is how the fact that. "'cow'
> refers to cows" is not absolutely unrevisable manifests
> itself - relative to the theory, "'cow' refers to cows"
> is a logical truth.   42

We might try to cash out the claim that terms in a mature science typically refer in terms of the relative infrequency with which we "scrap" such notions as 'electron'. But it would still be unclear how reference, so construed, could serve as an explanatory notion. Indeed, the sort of

explanatory role that Boyd proposes for reference seems to require pre-
cisely the sense of 'refers' that the internal realist needs to eschew
in order to avoid the clutches of the "model-theoretic' argument.

A second matter that remains obscure in those of Putnam's writings
presently under consideration is how we are to interpret the internal
realist's talk of _truth_, given our rejection of the metaphysical realist
picture. It is tempting to attribute to Putnam a Peircean notion of
truth on the basis of his remark that part of the moral to be drawn
from the "model-theoretic" argument concerns the _point_ at which meta-
physical realism collapses:

> Metaphysical realism collapses just at the point at which
> it claims to be distinguishable from Peircean realism -
> i.e. from the claim that there is an ideal theory (I don't
> mean that even _that_ claim isn't problematical, but it
> is problematical in a different way).   43

On the Peircean account, to claim that a statement S is true is not
to claim that it corresponds to some state of THE WORLD, but, rather,
to claim that S will be preserved within the Ideal Theory. Truth, so
construed, is taken to be _internal_ with respect to the scientific enter-
prise viewed diachronically: what it is for a statement or theory to be
true is elucidated in terms of an "ideal" application of the practices
and procedures of science. Such a view would seem to fit well with the
idea of internal realism as a "scientific picture".

On the other hand, Putnam's talk of unspecified "problems" with
the Peircean account of truth should lead us to exercise caution in
attributing a similar account to Putnam himself. It is interesting, in
this connection, to consider a footnote to the paper "Models and Real-
ity". Putnam begins by countering a possible objection to an identific-

ation of a statements being <u>true</u> with its "being verified, or accepted, or accepted in the long run:"

> It may be objected that a person could reasonably, and
> possibly truly, make the assertion:
> A; but it could have been the case that A and our scien-
> tific development differ in such a way to make A part of
> the ideal theory accepted in the long run; in that cir-
> cumstance, it would have been the case that A but it
> was not true that A.
>
> This argument is fallacious, however, because the different
> "scientific development" means here the choice of a
> different version; we cannot assume the <u>sentence</u> 'A'
> has a fixed meaning independently of what version we
> accept.   44

This is a curious passage. The objection seems to be an attempt to run something like Putnam's "modal" argument of the <u>Locke Lectures</u> against the Peircean account of truth, which eludes the argument in its original form. But Putnam's response, which seems to flirt with a Kuhnian view of sentence meaning, threatens his own later contention - especially in answer to certain criticisms by Hartry Field[45] - that different scien- tific versions can be genuinely <u>incompatible</u> with respect to their theoretical claims. A more obvious response to the objection is surely that it rests upon two assumptions that the Peircean need not accept: (a) the assumption that A's being the case, as asserted at the beginning of the objector's argument, is independent of the identity of the ideal theory, and, more importantly, (b) the assumption that the identity of the ideal theory depends on the vagaries of our actual scientific dev- elopment, so that different historical turns in science will yield <u>different</u> "Ideal Theories", rather than being an idealisation from scien- tific practice that is not vulnerable to the whims of historical devel- opment.

Of greater interest in the present context, however, are Putnam's

observations following the passage quoted above. He cites Dummett as a

fellow-subscriber to the view that

> what is involved is not that we identify truth with
> acceptability in the long run (is there a fact of the
> matter about what would be accepted in the long run?)
> but that we distinguish two truth-related notions: the
> internal notion of truth ("snow is white" is true if and
> only if snow is white), which can be introduced into any
> theory at all, but which does not explain how the theory
> is understood (because "snow is white" is true is under-
> stood as meaning that snow is white and not vice versa),
> and the notion of verification, no longer thought of as
> a mere index of some theory-independent kind of truth,
> but as the very thing in terms of which we understand
> the language..   46

The view of truth contained in this passage may remind us of the

view of reference encountered above in Putnam's defence of internal

realism against the charge that it is vulnerable to the "model-theo-

retic" argument. Both views seem to be deflationary in effect, a

therefore difficult to reconcile with the essentially reflationary

thrust of the "scientific picture" developed in the Locke Lectures

and "Reference and Understanding". I shall refrain from further spec-

ulation on these matters, however, and turn, with a certain measure

of relief, to Putnam's further exploration of our themes in Reason,

Truth, and History and related writings.

Endnotes - Chapter Two

1. Dummett, 1976c.

2. Ibid., p. 220.

3. Ibid., p. 221.

4. Ibid., p. 223.

5. Ibid., p. 224.

6. See especially Dummett, 1976a.

7. Dummett, 1976c, p. 225.

8. Putnam, 1976c.

9. Ibid., p. 226.

10. Loc. cit.

11. Loc. cit.

12. Loc. cit.

13. Loc. cit.

14. Putnam, 1978, pp. 5-6.

15. See, for example, Devitt, 1983a, and Demopoulos, 1982.

16. Putnam, 1976d, p. 123.

17. Ibid., p. 129.

18. Ibid., p. 132.

19. Devitt, 1983a, p. 296.

20. Putnam, 1976d, pp. 124-125.

21. Putnam, 1983a, pp. ix-x.

22. Putnam, 1976d, p. 125.

23. Ibid., pp. 125-126.

24. Ibid., p. 126.

25. Putnam, 1980, p. 475.

26. Putnam, 1976d, p. 129.

27. Putnam, 1980, p. 479.

28. Putnam, 1976d, pp. 128-129.

29. Ibid., p. 128.

30. Demopoulos, 1982.

31. Putnam, 1980, p. 480.

32. Ibid., p. 479.

33. Ibid., p. 479.

34. Ibid., pp. 480-481.

35. Ibid., p. 482.

36. Putnam, 1976d, p. 135.

37. Ibid., p. 136.

38. See, e.g., Quine, 1960, Chapter 2.

39. Putnam, 1976d, p. 136.

40. Ibid., p. 137.

41. Putnam, 1976a, p. 31.

42. Putnam, 1976d, p. 136.

43. Ibid., p. 130.

44. Putnam, 1980, p. 480.

45. Field, 1982.

46. Putnam, 1980, p. 480.

Chapter Three   "Externalist" and "Internalist" Perspectives

III-a   The Argument of Reason, Truth, and History

I have already had occasion to remark Putnam's observation, in the first of his Howison Lectures, that the principal task confronting the philosopher who follows Nelson Goodman in rejecting THE WORLD as a philosophical resource is to find "a narrow path between the swamps of metaphysics and the quicksands of cultural relativism and historicism." It is this task that Putnam has himself assayed in Reason, Truth, and History and other recent writings most of which are collected in the third volume of his Philosophical Papers.

·The argument of Reason, Truth, and History (henceforth RTH) turns upon a distinction, drawn at the beginning of the third chapter, between what Putnam characterises as "two philosophical perspectives", the "externalist" and "internalist" points of view. The former point of view is identified with metaphysical realism, and it is to be distinguished from the internalist perspective in the following respects: (a) for the externalist, "the world consists of some fixed totality of mind-independent objects": for the internalist, one can sensibly ask "what objects does the world consist of?" only from within a particular theory or description; (b) for the externalist, "there is exactly one true and complete description of 'the way the world is'": for the internalist, one may (though one need not) hold that there is more than one such "true" theory or description of the world; and (c) for the externalist, "truth involves some sort of correspondence relation between words or thought-signs and external things or sets of things": for the internalist, truth is

> some sort of (idealised) rational acceptability -. some
> sort of ideal coherence of our beliefs with each other
> and with our experiences <u>as those experiences are them-
> selves represented in our belief system</u> - and not corres-
> pondence with mind-independent or discourse-independent
> "states of affairs". There is no God's Eye point of view
> that we can know or usefully imagine; there are only the
> various points of view of actual persons reflecting
> various interests and purposes that their descriptions
> and theories subserve.

Chapters 1, 2, and 4 of <u>RTH</u> contain a sustained assault upon the externalist perspective, much of this being a reiteration or refinement of those arguments against metaphysical realism surveyed in the previous chapter of the present work. In arguing <u>for the</u> internalist perspective, which analyses truth in terms of rational acceptability, Putnam embarks upon a lengthy inquiry into the nature of rationality, an inquiry which provides the focus for the remaining chapters of the book. Theory of rationality is obviously of crucial importance from an internalist per- spective, for one's view of rationality will be directly reflected in one's theory of truth. For example, an internalist who held that there are no higher standards of what it is rational to believe than those provided by individual subjectivity or by cultural concensus would seem to be committed, thereby, to a relativistic view of truth. Alternatively, one who, like Putnam, takes truth to be an <u>idealisation</u> of rational acc- eptability, and who also defines rationality in terms of the procedures and practices of science, will arrive at a roughly Peircean theory of truth. In arguing for his own conception of truth, which purports to be both non-relativistic and pluralist, Putnam is accordingly led to offer arguments against relativistic and "scientistic" conceptions of ration- ality.

Underlying the discussion of the relative merits of the two "phil-

osophical perspectives", however, there is a perceived challenging of
more deeply entrenched philosophical distinctions - distinctions between
"objective" and "subjective" theories of truth, and, derivatively,
between "facts" and "values". It is these distinctions, according to
Putnam, that have distorted philosophical discussion of the nature of
truth and of rationality. It is assumed, for example, that if one gives
up the "objective" externalist notion of truth, then the only alternat-
ive is "the denial of the objectivity of truth and a capitulation to
the idea that all schemes of thought and all points of view are hope-
lessly subjective."[2] Or, again, it is assumed that, since the proced-
ures whereby we proceed in the sciences fail to yield anything in the
nature of normative ethical "facts", ethical beliefs can be no more
than reflections of "subjective" "values", and are therefore lacking
in cognitive significance. Or, finally, it is assumed that one can
rationally deliberate only about means, which are instrumental and
"objective", but not about ends, which are ultimately "subjective".

Putnam contends that working within these sorts of dichotomies
has engendered "alienated" views of truth and rationality which "cause
one to lose one part or another of one's self and the world, to see the
world as simply consisting of elementary particles swerving in the void,
...or to deny that there is a world at all, as opposed to a bunch of
stories that we make up for various (mainly unconscious) reasons."[3]
The internalist perspective is offered as a framework within which we
may develop "a non-alienated view" comprising a "Kantian" conception
of truth that "unites subjective and objective components." Putnam
maintains that the internalist can preserve the "realist intuition" that
"a statement can be rationally acceptable at a time but not true."

,while also preserving the "subjectivist" insight that "there is no fixed ahistorical organon which defines what it is to be rational."[4] He resists the relativist conclusion that there are no constraints on what can be rational, however, arguing that "our notion of rationality is, at bottom, just one part of our conception of human flourishing, our idea of the good."[5]

In filling in some of the details of this picture, I shall limit my attention to those matters which have an immediate bearing on the project in hand. I shall examine only those considerations which Putnam brings forward against the externalist perspective, and the manner in which, so he contends, the internalist is able to avoid such pitfalls of externalism. In this context, I shall also refer to certain later writings that bear upon these themes. I shall not, however, further elucidate Putnam's critique of relativistic and "scientistic" theories of rationality, nor shall I discourse upon the positive aspects of his programme - his own theory of rationality and the internalist notion of truth grounded thereupon. The reader will find a few brief allusions to such matters in the concluding chapter of this work, but further discussion falls beyond the scope of the present project.

III-b   The Infirmities of Externalism

1) Brains in a Vat

In the opening two chapters of RTH, Putnam develops in greater detail his charge that externalism (metaphysical realism) is unable to fund the sort of determinate relation of reference required by the ex-ternalist account of truth. The argument, as before, proceeds from within the externalist framework, and is premissed on the assumption that "words stand in some sort of one-one relation to (discourse-independent) things and sets of things."[6] It is argued, firstly, that we can make no sense, from within the externalist framework, of the claim - often made by externalists in the interests of illustrating the contention that even an "Ideal Theory" might be false - that "we might be permanently deluded brains in a vat." It is then argued that the externalist framework lacks the resources to make sense of the claim that there is a unique, determinate reference relation between words and things in THE WORLD.

Putnam's discussion of the "brain in a vat" hypothesis presupposes two assumptions about reference defended in the opening pages of RTH. Putnam argues that (a) we should reject any "magical" theory of refer-ence which postulates some sort of "necessary connection between a rep-resentation, mental or otherwise, and what it represents", and that (b) "one cannot refer to certain kinds of things, e.g. trees, if one has no causal interaction at all with them."[7] His arguments draw upon a number of hypothetical examples. In arguing for the first assumption, he offers such examples to illustrate that one can have individuals whose relevant mental or psychological states are identical, yet whose language differs in its reference. Thus, he maintains, "even a large

and complex system of representations, both verbal and visual, still does not have an <u>intrinsic</u>, built-in, magical connection with what it represents - a connection independent of how it was caused and what the dispositions of the speaker or thinker are."[8] This conclusion is not to be evaded by maintaining that reference is secured not by mental images but by the possession of "concepts", construed as mental objects that necessarily refer: for, so Putnam argues, "concepts" are not mental objects at all, but, rather, <u>abilities</u> to <u>use</u> language in certain ways. Thus we should reject the idea that reference is fixed by "what goes on inside our heads". This claim is supported by "Twin Earth" examples similar to those offered in "The Meaning of 'Meaning'". We hypothesise that speakers on Earth and Twin Earth have mental states that are qualitatively identical in all respects that might be relevant to fixing the reference of the term 'water'. But if the physical "stuff" causally related to the use of the term differs - if it is $H_2\theta$ on Earth and XYZ on Twin Earth - then, so Putnam argues, the <u>reference</u> of the term also differs: "The mental state by itself, in isolation from the whole situation, does not fix the reference."[9] He further argues that even the ability to <u>use</u> a language in accordance with the relevant communal norms may be insufficient to secure referential solidarity with the linguistic community. An individual might pass a Turing-type test of competence, yet fail to refer as other speakers do because of a failure to engage in the sorts of "non-verbal transactions" with extra-linguistic reality that are necessary if reference is to be established: "Even if we consider not words by themselves but...<u>programs</u> <u>for using words</u>...unless those programs themselves <u>refer to something</u> <u>extralinguistic</u> there is still no determinate reference that those words

possess."[10]

We are asked to consider a particular version of the "brain in a vat" hypothesis. It is supposed that all sentient beings are permanently isolated brains in a given vat, programmed to experience a hallucination qualitatively identical to our own (supposed!) experience of the world. We are also to suppose that there is no "evil genius" outside the vat responsible for this state of affairs - "the universe just happens to consist of automatic machinery tending a vat full of brains and nervous systems" - and that there never _was_ such an intelligent creator-designer. Finally, it is assumed that "the automatic machinery is programmed to give us all a _collective_ hallucination, rather than a number of separate unrelated hallucinations. Thus, when I seem to myself to be talking to you, you seem to yourself to be hearing my words."[11] Putnam then poses the following question:

> Suppose this whole story were actually true. Could we,
> if we were brains in a vat in this way, _say_ or _think_
> that we were?   [12]

His contention is that the hypothesis that we are all brains in a vat in the envisaged circumstances cannot possibly be true because it is "self-refuting" in the sense that "if we can consider whether it is true or false, then it is not true."

This conclusion seems unacceptable because there is surely a "physically possible world" in which all sentient creatures are envatted in the prescribed manner, and it would seem that if _we_, who, _ex hypothesi_ share all experiences (qualitatively speaking) with the denizens of such a possible world, can give an argument to show that we are not in such a situation, then so could _they_. Putnam responds that

> although the people in that possible world can think and
> "say" any words we can think and say, they cannot (I
> claim) <u>refer</u> to what we can refer to. In particular,
> they cannot think or say that they are brains in a vat
> (even by thinking "we are brains in a vat"). 13

For, given our assumptions about reference, we cannot infer, from the

fact that envatted and unenvatted brains share all mental states, qual-

itatively construed, that they therefore refer to the same things by

means of their visual and verbal representations. In fact, when an en-

vatted brain thinks, e.g. "There is a tree in front of me", "his

thought does not refer to actual trees...[but] to trees in the image,

or to the electronic impulses that cause tree-experiences, or to the

features of the program that are responsible for those electronic im-

pulses."[14] Since this will also apply to the use of such terms as 'vat'·

and 'nutrient fluid' in "vat-English",

> it follows that if their "possible world" is really the
> actual one, and we are really the brains in a vat, then
> what we now mean by "we are brains in a vat" is that
> <u>we are brains in a vat in the image</u> or something of that
> kind (if we mean anything at all). But part of the hypo-
> thesis that we are brains in a vat is that we aren't brains
> in a vat in the image (i.e. what we are "hallucinating"
> isn't that we are brains in a vat). So, if we are brains
> in a vat, then the sentence "We are brains in a vat",
> says something false (if it says anything). In short, if
> we are brains in a vat, then "We are brains in a vat"
> is false. So it is (necessarily) false. 15

One moral to be drawn from this, according to Putnam, is that we

should not take "physical possibility" - compatibility with the laws of

physics.- as the touchstone of "what might actually be the case". What

rules out the possibility that we are brains in a vat is "not physics

but philosophy."[16] The philosophical strategy employed here is akin to

the Kantian notion of a transcendental investigation, Putnam maintains,
in that it proceeds by inquiring as to "the preconditions for thinking
about, representing, referring to, etc."; it involves reasoning a priori
in that it investigates "what is reasonably possible assuming certain
general premisses, or making certain very broad theoretical assump-
tions."[17]

2) The Problem of Reference

As a number of commentators have observed,[18] Putnam's analysis of the "brain in a vat" hypothesis does not seem to pose too serious a challenge to the externalist. He may indeed have to relinquish, or at least reformulate, a favourite model of realist truth, but the claim that there is a determinate reference relation between language and entities in THE WORLD appears to survive untarnished. Indeed, it is to this very notion of a unique "intended" interpretation of our language that Putnam himself implicitly appeals when he contends that the terms in vat-English should be assigned specific referents "in the image rather than referents outside the vat identical to the referents of homophonic terms in ordinary English. The externalist's traditional contention that we _might_ all be brains in a vat can be blamed upon an uncritical acceptance of the thesis that reference is somehow determined by the mental-state of language users. Suitably chastened by Putnam's argument, the externalist can acknowledge the implications of the causal constraints on reference and then proceed much as before. In chapter 2 of _RTH_, however, Putnam argues that the externalist account of reference is more radically flawed. The arguments are essentially variations upon the "model-theoretic" theme: it is claimed that the unique "intended" interpretation of our language in THE WORLD cannot be picked out by either operational and theoretical constraints, or our referential intentions, or a physicalistic theory of reference.

Putnam describes the idea that the interpretation of a natural language is fixed by the collective agency of "operational and theoretical constraints" as "the received view". On more sophisticated versions of the received view, reference is a function not of our _current_

"operational and theoretical constraints", but of those constraints that would be accepted under ideal conditions of inquiry. Putnam claims that even the "sophisticated" version of the received view fails because "it tries to fix the intensions and extensions of individual terms by fixing the truth-conditions for whole sentences."[19] In the case of the "sophisticated" version, this is equivalent to the claim that even an "ideal theory" - one which satisfied ideal operational and theoretical constraints - would admit of a plurality of "acceptable" interpretations. Putnam argues that, even if we specify the truth-value of every sentence <u>in every possible world</u>, the received view will not work: "There are always infinitely many different interpretations of the predicates of a language which assign the 'correct' truth-values to the sentences in all possible worlds, <u>no matter how these 'correct' truth-values are singled out.</u>"[20] In an Appendix, Putnam offers a technical proof of this result, and in the body of the text he illustrates the method of the proof for the case of a single sentence, "A cat is on a mat", henceforth (1). In the standard interpretation of (1), 'cat' refers to cats and 'mat' refers to mats. Putnam then offers a reinterpretation under which (1) comes to mean "A cat* is on a mat*", where 'cat*' and 'mat*' receive complex definitions in terms of the standard interpretation of the language. The reinterpretation has the effect that "in the actual world 'cat' refers to <u>cherries</u> and 'mat' refers to <u>trees</u> without affecting the truth-value of (1) in any possible world."[21] He further argues that one cannot reject this non-standard interpretation on the grounds that cat* and mat* are "queer" or merely "extrinsic" properties, because what one counts as a "queer" or "extrinsic" property depends upon how one interprets <u>other</u> predicates, or upon which

sets of properties one takes as basic. (The discussion, here, is reminiscent of Goodman's defence of the predicates 'grue' and 'bleen'[22].)

In the remaining pages of Chapter 2, Putnam canvasses and rejects a number of alternative externalist suggestions as to how terms acquire a determinate reference. To the proposal that evolution, and the requirements for human survival, have produced a unique correspondence between our representations and external things, he responds that, while survival does presuppose a tendency to have mainly true beliefs, true beliefs do not, as already argued, serve to single out a unique relation of correspondence. Nor can our intentions to refer to certain external things rather than other external things fix the reference of our terms; for one can intend to refer to something only if one already has some conceptual machinery that enables one to pick out that thing - i.e. only if one can already refer to that thing. Finally, Putnam returns to Field's contention that reference is a "physicalistic" relation. If this is construed as the claim that we can define 'reference' in terms of some physical relation R, then the response is that such a definition, like a "causal theory" of reference, is "just more theory" which itself requires interpretation. The proposed definition will be true in each admissible model of the language, but it will not serve to reduce the number of admissible models. If, on the other hand, it is insisted that it is R itself, as a relation in THE WORLD, that determines the reference of our language, including the language used in giving the definition of 'reference', then we need to explain (a) how we could ever learn to express this thesis -i.e. how we could ever acquire the ability to refer to R, and (b) what it is that makes it true that R is the relation of reference, given the plethora of other "correspondence" relat-

ions between language and THE WORLD:

> It seems as if the fact that R is reference must be a
> metaphysically unexplainable fact, a kind of primitive,
> surd, metaphysical truth...To me, believing that some
> correspondence just is reference (not as a result of our
> operational and theoretical constraints, or our inten-
> tions, but as an ultimate metaphysical fact) amounts to
> a magical theory of reference...Even if one is willing
> to contemplate such unexplainable metaphysical facts,
> the epistemological problems that accompany such a meta-
> physical view seem insuperable. For, assuming a world
> of mind-independent, discourse-independent entities...,
> there are, as we have seen, many different "corresponden-
> ces" which represent possible or candidate reference
> relations (infinitely many, in fact, if there are infin-
> itely many things in the universe)...But then there are
> infinitely many different possible "surd metaphysical
> truths" of the form "R is the real (metaphysically
> singled-out) relation of reference"...Note that all
> these infinitely many metaphysical theories are com-
> patible with the same sentences being true, the same
> "theory of the world", and the same optimal methodology
> for discovering what it true.   23

A further rehearsal of the "model-theoretic" considerations occurs

near the end of chapter 3 of RTH, where Putnam considers the suggestion

that truth be defined in terms of "some kind of abstract isomorphism,

or, if not literally an isomorphism, some kind of abstract mapping of

concepts onto things in the (mind-independent) world.[24] The problem

with such a suggestion, Putnam avers, is that there are too many such

"correspondences" between words and other entities, and one could single

out one of these "correspondences" as the relation of reference only if

one already had access to the "noumenal" objects. Putnam supports this

claim by offering an example of theories that are metaphysically "incom-

patible" yet intertranslatable. If Newtonian physics were true, then we

could give a description of every physical event either in terms of

Maxwell's field equations or in terms of action-at-a-distance and "retar-

ded potentials". The realist would have to regard the two resultant Newtonian theories as metaphysically incompatible, since "either there are or there aren't causal agencies (the fields) which mediate the action of separated particles on each other":

> But the two theories are mathematically intertranslatable. So if there is a "correspondence" to the noumenal things which makes one of them true, then one can define another correspondence which makes the other theory true. If all it takes to make a theory true is abstract correspondence (never mind which), then incompatible theories can be true. 25

This conclusion is unacceptable to the metaphysical realist, for whom there has to be a fact of the matter as to <u>which</u> of two such incompatible theories "gets THE WORLD right". At this point, Putnam cites the familiar "model-theoretic" considerations as an indication that the realist faces even more serious difficulties:

> Not only may there be correspondence between objects and (what we take to be) incompatible theories (i.e. <u>the same objects</u> can be what logicians call a "model" for incompatible theories), but even if we fix the theory <u>and</u> fix the objects, there are (if the number of objects is infinite) infinitely many <u>different</u> ways in which the same objects can be used to make a model for a given theory. This simply states in mathematical language the intuitive fact that to single out a correspondence between two domains one needs some independent access to both domains. 26

A related infirmity of the externalist picture, according to Putnam, is an implicit commitment to the thesis that THE WORLD contains "Self-Identifying Objects". When we use a term like 'horse', the extension of the term includes not only those horses with which we have causally interacted, but also all other things "of the same kind". But "same kind" has as many interpretations as there are categorial systems

2
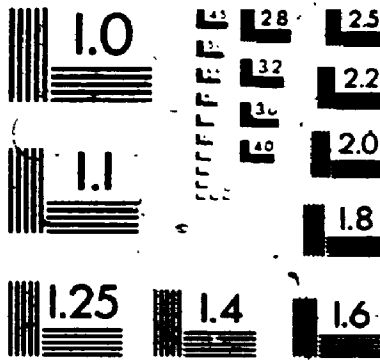
classifying certain properties as relevant, and other properties as
irrelevant, to "sameness of kind". The externalist, however, is comm-
itted to the view that there is one such categorial system that is
"objectively right" in that it corresponds to the way in which THE
WORLD itself "sorts things into kinds". This thesis - that there are
objects that are of the same kind _in themselves_ - is the thesis that
the world contains "Self-Identifying Objects".[27]

3) Externalism and Skepticism

Putnam develops a different kind of criticism of externalism in chapter 4 of RTH. He argues that externalism engenders various irresoluble skeptical puzzles - "irresoluble" within the externalist framework, that is. He further argues, as we shall see, that no such problems arise from an internalist perspective, and it is clear that he does not see this as an intellectual loss. This line of reasoning is in fact prefigured in chapter 3, where we are asked to consider whether one might always be wrong about one's past sensations. Suppose, for example, that it is hypothesised that what we term a 'sensation of red' at even-numbered minutes is qualitatively similar to what we call a 'sensation of green' at odd-numbered minutes, but that our memories systematically deceive us and we never notice this fact. The externalist must hold that we might indeed fall prey to such an error, Putnam claims. For the externalist

> my previous sensations either are or aren't similar to
> the sensations I now describe by the various verbal labels···
> 'sensation of red', 'pain', etc., and whether they are
> or aren't is a totally different question from whether I
> then classified them under those same verbal labels.   · 28

Putnam suggests that there is something questionable in the claim that our manner of clasifying our sensations, in such a case, is "really wrong" when, ex hypothesi, we never notice what is happening and must therefore be completely successful in those of our actions which are predicated upon such a classification.

The skeptical difficulties canvassed in chapter 4 concern the relationship that obtains between physical states and mental states. In the most fully developed example, Putnam considers the implications

of combining an externalist view of truth with the mind-body "iden-
tity theory". If the identity theory is correct, then a given sensation
state Q will be synthetically identical with some brain state B. The
externalist must hold that there is a fact of the matter, independent
of our choice or our linguistic conventions, as to precisely which
brain state is synthetically identical with Q. But, according to the
current theories in cognitive psychology, the occurrence of, e.g. a
particular visual sensation, is correlated with a number of distinct
neurological events occurring in different regions of the brain. In
"split-brain" patients, some of these events may occur in the absence
of others. With which of these neurological events, then, is the occur-
rence of the visual sensation to be identified? This is not a question
that can be answered empirically, Putnam maintains: "The difficulty is
that there are identity theories which are observationally indistin-
guishable, by which I mean that they lead to the same predictions with
respect to the experiences of all observers in normal conditions."[29]
But the externalist is committed to there being a fact of the matter,
as to which of these theories is correct. An analogous problem arises
when we consider Thomas Nagel's claim that we cannot imagine what it
would be like to be a bat. We can formulate two theories, according to
which bat qualia are taken to be qualitatively identical to, and qual-
itatively distinct from, human qualia, respectively, and these two
theories will generate "the same predictions with respect to what human
observers, normal and abnormal, will experience." Again, the external
ist must hold that there is a fact of the matter as to which theory is
correct.

Finally, Putnam presents three further theories which are surely

untrue, but which the externalist must take to represent genuine poss-
ibilities. One of these theories hypothesises that rocks have qualia -
i.e. that events qualitatively similar to human sensations take place
in rocks. The externalist can regard such a hypothesis as highly im-
probable, but he must allow that it is logically possible because this
follows from his view of THE WORLD, as containing "Self-Identifying
Objects", and his view of reference, as a way of somehow grasping such
objects:

> The metaphysical realist....thinks that [my having sensat-
> ion Q at time t and some physical event in a rock] might
> (logically possibly) be [qualitatively similar], even
> though it is "crazy" to think so. And he thinks this
> because he is under the illusion that by having the sen-
> sation in question, with its qualitative character, its
> "the way it feels", with its functional role, with the
> accompanying thoughts and judgments, he has somehow
> brought it about that the expression, "the way this
> sensation feels" (or some technical substitute...)
> refers to one definite "universal", one absolutely
> well-defined property of metaphyscial individual events.
> But this is not the case.   30

4) Later Reflections

Realism and Reason, the third volume of Putnam's Philosophical
Papers, contains a number of interesting clarifications, augmentations,
and defences of the case against externalism developed in RTH and ear-
lier writings. The most extensive discussion occurs in the paper, "Why
There Isn't a Ready-Made World", the first of the 1981 Howison Lectures.
In this paper, Putnam doesn't explicitly address the "model-theoretic"
argument, but focuses, rather, upon what he terms "materialism", as
the predominant contemporary form of metaphysical realism. More specif-
ically, he considers a materialist strategy for evading the charge that
the metaphysical realist cannot specify constraints that might single
out one among a number of correspondences as the unique relation of
reference. The materialist, as a metaphysical realist, is committed to
there being "One True Theory", and thus, in so far as he also espouses
a correspondence theory of truth, to the existence of a "Ready-Made
World": "The world itself has to have a 'built-in structure' since
otherwise theories with different structures might correctly 'copy'
the world (from different perspectives) and truth would lose its abso-
lute (non-perspectival) character."[31] Causal relations, the materialist
might claim, are part of the "built-in" structure of THE WORLD, and
reference is itself part of the network of such causal relations. It
is thus certain facts about THE WORLD that single out a particular
correspondence as the unique relation of reference.

Putnam responds that a materialist, who views the denizens of THE
WORLD as mind-independent entities ultimately describable solely in
terms of their physical properties (properties definable in terms of the
"fundamental magnitudes" of current or future physics), cannot consis-

tently hold that either "causal relations" in general or "reference" in particular are "<u>physical</u> relations" so construed. I defer consideration of the details of this argument to chapter five below.

In the concluding paragraphs of the first <u>Howison Lecture</u>, Putnam comments further on the metaphysical materialist's project, which he characterises as the quest for "a coherent theory of the noumena: consistent, systematic, and arrived at by 'the scientific method'". Such a project is chimerical, he claims, and alien to the very spirit of science. For, as he has argued in earlier writings, such paradigms of scientific achievement as Newton's theory of gravity or the physics that has replaced it "admit of a bewildering variety of empirically equivalent formulations." And, he further maintains, scientists themselves take no interest in speculations as to which of the different "metaphysical interpretations" of physical theory is correct:

> There is, then, nothing in the history of <u>science</u> to
> suggest that it either aims at or should aim at one
> single <u>absolute</u> version of "the world". On the contrary,
> such an aim, which would require science itself to de-
> cide which of the empirically equivalent successful
> theories in any given context was "really true", is
> contrary to the whole spirit of an enterprise whose
> strategy from the first has been to confine itself to
> claims with clear <u>empirical</u> significance.   32

In what sense can one remain a "scientific realist" while espousing such a view? This issue is clarified in a paper, "Three Kinds of Scientific Realism"[33], which appeared approximately a year after the delivery of the <u>Howison Lectures</u>. Putnam rejects two conceptions of "scientific realism" which partake in the "metaphysical materialism" criticised in the first <u>Howison Lecture</u>. One conception, characterised as "scientific imperialism - physicalism, materialism", cannot give an adequate account

of such properties as truth, reference, and justification. The second
conception, metaphysical realism with respect to science, involves the
"unacceptable" thesis that there is a "fact of the matter" as to which
of two mathematically and empirically equivalent theories is correct.
Putnam then identifies and endorses a third conception of "scientific
realism" which involves an adherence to two principles. Firstly, one is
to take theoretical discourse in science "at face value": "Electrons
exist in every sense in which chairs (or sensations) exist; electron
talk is no more derived talk <u>about</u> sensations or 'observable things'
than talk about sensations or chairs is derived talk about electrons."[34]
Secondly, one should retain "the ideal of scientific convergence",
which is "a very central part of the scientific outlook - a part which
informs scientific methodology in a host of ways." To retain this ideal
is to apply the Principle of Charity, "the principle that we <u>should</u>
often identify the referents of terms in different theories so as to
avoid imputing too many false or unreasonable beliefs to those we are
interpreting."[35]

    <u>Realism and Reason</u> contains two further interesting contributions
to Putnam's critique of externalism that deserve mention: firstly, the
diagnosis of an additional problem attending the externalist commitment
to a "ready-made world", and, secondly, a clarification of Putnam's
reasons for dismissing a particular line of response to the "model-
theoretic" argument.

    In the paper, "Vagueness and Alternative Logic", Putnam argues
that the externalist has a problem giving an adequate semantics for
sentences containing <u>vague</u> expressions. (One such sentence, cited by
Dummett in his Jerusalem Conference response to Putnam, is "The number

of trees in Canada is even.") The externalist is committed to the
following framework:

> Assume a definite totality T of all states of affairs.
> Assume a relation R (the metaphysically singled-out
> relation of correspondence) such that a sentence is true
> if and only if the sentence (as used on the particular
> occasion) corresponds under R to a member of the total-
> ity T which actually obtains. States of affairs are, of
> course, to be thought of as non-mental, non-linguistic
> entities which determinately obtain or do not obtain
> no matter what we think or say.   36

This yields the following definition of 'truth':

$$S \text{ is } \underline{true} \text{ iff } (\exists X)((X \text{ belongs to } T) . (R\{S, X\}) . (X \text{ obtains}))$$

The problem with vague sentences arises because all of the terms that
occur in the expression on the right of 'iff' in the foregoing defini-
tion must be non-vague for the externalist: for both states of affairs
and relations between sentences and states of affairs must themselves
belong to the realm of "discourse-independent objects and properties",
and the externalist perspective does not allow for the possibility of
vague objects. The externalist definition of 'truth' therefore implies
that "there is a 'fact of the matter' as to whether a sentence S is
true or not, which is just what sometimes fails to be the case with
vague sentences!"[37]

Putnam considers a common externalist response to this sort of
reasoning, namely, that we can elucidate the truth of vague sentences
by appeal to an "ideal language" which would "rationally reconstruct"
ordinary discourse. We can say that "sentences in ordinary language are
true or false relative to translations (or "reconstructions") into the
ideal language." But this strategy fails, he claims:

> If the translation is "reasonable", is <u>this notion</u>
> <u>"reasonable" itself a notion in the ideal language or</u>
> <u>only a notion in ordinary language?</u> If the latter, then
> we have answered the challenge, "In what sense can a
> sentence which employs vague notions be <u>true</u>?", with
> an answer which uses a vague notion. [38]

The internalist, on the other hand, has no such difficulties with

vague language, Putnam maintains. Not only can he embrace vague pred-

icates to the extent that the latter serve our linguistic purposes,

but also "since I don't think that <u>any</u> objects are totally mind-inde-

pendent (or theory-independent), I would add that, on my view, <u>objects</u>

and <u>properties</u> are in general vague too."[39]

Turning to the second point, in the Introduction to <u>Realism and</u>

<u>Reason</u>, Putnam sets out a helpfully simplified version of the "model-

theoretic" argument and an equally helpful classification of common

responses to that argument. We are offered a parable in which God,

moved by boredom, plays with the "correspondences" between language and

THE WORLD:

> Imagine that C_1 and C_2 are admissible "correspondences"
> (satisfaction relations), i.e. that C_1 (respectively
> C_2) is the satisfaction that one gets if M_1 (respectively
> M_2) is the model that one uses to interpret English,
> where M_1 and M_2 are both models which satisfy all the
> operational and theoretical constraints that our prac-
> tice imposes. Then what He did (to Hebrew, to Assyrian,
> to Coptic...) was to specify that when a <u>man</u> used a
> word, the word would stand for its image or images
> under the correspondence C_1, and that when a <u>woman</u>
> used a word, the word would stand for its image or
> images under the correspondence C_2...Notice that the
> same sentences are true under both of His reference-
> assignments, the sentences we accept generate the same
> experiential expectations under both schemes, the beha-
> viour that is associated with believing-true or desiring-
> true particular sentences is the same under both schemes,
> and if the expectations we have or the things we do are
> successful (respectively, unsuccessful) the sentences
> we are <u>then</u> required to accept by our operational and

> theoretical criteria are the same and their truth-
> values are the same.   40

The point of this "parable" is to give a more intuitive rendering of
the "model-theoretic" result that no amount of "operational and theoret-
ical constraints" imposed by our practice can determine a unique "corr-
espondence" between language and THE WORLD. A common externalist res-
ponse to the "model-theoretic" argument, however, is to argue that
there are further constraints, available to the externalist, which do
serve to fix reference in the required fashion. These further con-
straints are provided by "causal relations", or "structure in THE WORLD";
or something of that sort, the externalist claims. I have already ref-
erred to Putnam's discussion of such strategies in the first of the
Howison Lectures. In the Introduction to Realism and Reason, he brings
out nicely various stages in the debate between the externalist and the
internalist on this matter.

The debate begins with the externalist's attempt to counter the
"model-theoretic" argument by appealing to a "causal theory" of ref-
erence. In learning a language, it is said, a child associates terms
with certain "mental representations" which are themselves caused by
certain external objects or events. It is the latter that are the ref-
erents of the terms as employed by the language-learner. The internalist
responds to this by arguing that a "causal theory" of reference is
"just more theory" which itself admits of as many different interpret-
ations as the rest of our body of beliefs. In the context of Putnam's
"parable", this amounts to pointing out that the appeal to a "causal
theory" of reference could be made by both a male and a female philos-

opher. Each would be equally "correct", but this would in no way under-
mine God's ability to select different "correspondence" relations for
males and females as His whims dictate.

The second stage of the debate begins with the externalist char-
ging obfuscation. The realist position, it is claimed, is not that the
"causal theory", qua linguistic object, provides the additional con-
straint on reference, but that reference itself, as a causal relation
in the world, provides the necessary constraint on the interpretation
of the language, including the language employed in formulating the
"causal theory" of reference. I shall examine various versions of this
externalist contention in chapter five below. For the present, I shall
simply note Putnam's general strategy against such a contention. He
argues that the externalist is guilty of "ignoring his own epistemolog-
ical position" in assuming that his own use of the terms 'cause' and
'refer' uniquely refers to a particular relation in THE WORLD. But the
question at issue is precisely how any term can "uniquely refer".

Finally, the more sophisticated externalist may argue that Putnam
has only shown that reference is not fixed by anything psychological.
It is therefore quite possible that reference is fixed by something
non-psychological. Putnam's response to this suggestion is that "to
think that a sign-relation is built into nature is to revert to medie-
val essentialism, to the idea that there are 'self-identifying objects'
and 'species' out there."[41]

III-c   The Internalist Perspective

At the beginning of chapter 3 of RTH, Putnam states that the sig-
nificance of the discussion of the "brain in a vat" hypothesis resides
chiefly in the fact that it brings out the difference between the exter-
nalist and internalist perspectives. For the externalist, the truth of
a theory consists in its corresponding to the the ~~way~~ THE WORLD is in
itself. Since the "brain in a vat" hypothesis represents a physical
possibility, and thus, it would seem, a way in which THE WORLD might
be, the externalist wishes to maintain (at least initially) that the
hypothesis might be true. For the internalist, on the other hand, the
truth of a theory consists in its fitting, not "THE WORLD as it is in
itself", but, rather, "the world as the world presents itself to some
observer or observers."[42] Truth, for the internalist, is a form of ideal
coherence between beliefs, and between beliefs and experiences "as those
experiences are themselves represented in our belief system."[43] Thus
there can be truth only where there are observers whose beliefs can
exhibit such coherence. But there are, ex hypothesi, no sentient beings
outside the "vat" for whom the "brain in a vat" hypothesis could ever
be rationally acceptable. The claim that the hypothesis might be true
thus presupposes that truth is "independent of observers altogether",
that is, it presupposes the externalist conception of truth. Thus the
internalist can dismiss the "brain in a vat" hypothesis as "only a
story, a mere linguistic construction, and not a possible world at
all."[44]

Putnam also reiterates the claim, first made in "Realism and
Reason", that, unlike the externalist, the internalist ("internal

realist") can easily explain how linguistic and mental representations can uniquely refer. In the present context, however, he places less emphasis upon the "contextual a priority" of reference-assigning statements, stressing, rather, the manner whereby "versions" and their objects come into being:

> In an internalist view..., signs do not intrinsically correspond to objects, independently of how those signs are employed and by whom. But a sign that is actually employed in a particular way by a particular community of users can correspond to particular objects within the conceptual scheme of those users. "Objects" do not exist independently of conceptual schemes. We cut the world up into objects when we introduce one or another scheme of description. Since the objects and the signs are alike internal to the scheme of description, it is possible to say what matches what. Indeed, it is trivial to say what any word refers to within the language the word belongs to, by using the word itself. What does 'rabbit' refer to? Why, to rabbits, of course. What does 'extraterrestrial' refer to? To extraterrestrials (if there are any)...For me, there is little to say about what reference is within a conceptual scheme other than these tautologies.   45

Clarifying the internalist perspective requires that it be distinguished not only from externalism but also from various other non-externalist options - and, in particular, from divers forms of relativism. The burden of making this latter distinction falls mainly upon the later chapters of RTH, but the groundwork is laid in chapter 3. Internalism, we are told, is not "a facile relativism", where the latter seems to be equated with the view that "every conceptual scheme is... just as good as every other." Putnam maintains that the inadequacies of such a view readily reveal themselves to one who selects, for example, a conceptual scheme according to which humans can engage in unassisted flight. Internalism, on the other hand, avoids such implications:

> Internalism does not deny that there are experiential
> inputs to knowledge; knowledge is not a story with no
> constraints except internal coherence; but it does deny
> that there are any inputs which are not themselves to
> some extent shaped by our concepts, by the vocabulary
> we use to report and describe them, or any inputs which
> admit of only one description, independent of all con-
> ceptual choices.   46

The internalist does not identify truth with rational acceptability,

where the latter is conceived as largely a matter of "coherence and

fit" between beliefs, both "theoretical" and experiential. Truth cannot

be so identified because "truth is supposed to be a property of a state-

ment that cannot be lost, whereas justification can be lost."[47] (This

is, of course, the "pre-critical" "modal" argument against "verificat-

ionism".) Rather, truth, for the internalist, is

> an idealization of rational acceptability. We speak
> as if there were such things as epistemically ideal
> conditions, and we call a statement "true" if it could
> be justified under such conditions...The two key ideas of
> an idealization theory of truth are (1) that truth is
> independent of justification here and now, but not in-
> dependent of all justification. To claim a statement
> is true is to claim it could be justified. (2) truth
> is expected to be stable or "convergent"; if both a
> statement and its negation could be "justified", even
> if conditions were as ideal as one could hope to make
> them, there is no sense in thinking of the statement
> as having a truth-value.   48

Putnam cites Kant's "Critical Philosophy" as a historical precur-

sor of internalism. He suggests that Kant should be read as proposing

what might be termed an "extended Lockean view"[49]. On such a view, we

apply Locke's account of "secondary qualities", as "properties - nature

unspecified - which enable the object to affect us in a certain way",

to all qualities. If we adopt such a view

> it follows that everything we say about an object is of
> the form: it is such as to affect us in such-and-such a
> way. Nothing at all we say about any object describes the
> object as it is "in itself", independently of its effect
> on us, on beings with our rational natures and our bio-
> logical constitutions...Our ideas of objects are not
> copies of mind-independent things. 50

Putnam's internalist reading of Kant rests upon two further claims con-
cerning Kant's purported "extended Lockean" views. Firstly, Putnam
maintains, Kant applies his analysis not only to external objects, but
also to sensations ("objects of inner sense"); this precludes retain-
ing a correspondence theory of truth for "inner events". Secondly, Kant
denies there is any "one-to-one correspondence between things-for-us
and things-in-themselves", e.g. a "noumenal" chair which "corresponds"
to our representation of a chair in the sense that the former possesses
those specific Powers which produce in us those qualities which char-
acterise the latter. Rather, "Powers" are to be ascribed to "the whole
noumenal world." The resultant conception of truth is akin to the inter-
nalist conception outlined above. A true statement "is a statement that
a rational being would accept on sufficient experience of the kind that
it is actually possible for beings of our nature to have...Truth is
ultimate goodness of fit."51

In the previous section, we noted that the externalist perspective
is purportedly beset with difficulties relating to (a) the possibility
of metaphysically incompatible but mathematically intertranslatable
theories, (b) the need to postulate "Self-Identifying Objects" in THE
WORLD, and (c) the postulation of unknowable facts that engender irresol-
uble skeptical disputes. Putnam contends that internalism is free from,
such difficulties:

(a) The internalist can quite happily embrace the possibility of true but incompatible theories, it is claimed: "Why should there not sometimes be equally coherent but incompatible conceptual schemes which fit our experiential beliefs equally well? If truth is not a (unique) correspondence, then the possibility of a certain pluralism is opened up."[52]

(b) The internalist can avail himself of a notion of "self-identifying objects" which is not available to the externalist and is not metaphysically problematic in the way that the externalist's notion is:

> If, as I maintain, "objects" themselves are as much made as discovered, as much products of our conceptual invention as of the "objective" factor in experience, the factor independent of our will, then of course objects intrinsically belong under certain labels; because those labels are the tools we used to construct a version of the world with such objects in the first palce. But this kind of "Self-Identifying Object" is not mind-independent. [53]

(c) The externalist's skeptical difficulties arise from his commitment to there being a fact of the matter - albeit unknowable - as to whether any two events are "qualitatively similar" or "synthetically identical", where such "facts" pertain to the way THE WORLD is independently of our epistemic access to such matters. For the internalist, on the other hand, "it does not follow that there is a fact of the matter in every case as to whether two sensations (let alone two arbitrary events) are qualitatively similar or dissimilar."[54] On an internalist account, the qualitative similarity or dissimilarity of two sensations depends not upon how THE WORLD itself sorts things into kinds, but upon which judgment concerning the sensations would be rationally acceptable under suitably ideal epistemic conditons. And this may depend crucially upon the concepts that it suits our purposes to employ, and the overall

coherence of our system of beliefs. Putnam urges that we view such issues as the "correct" formulation of the "identity theory", or the "correct" answer to the question about the _qualia_ of bats, as matters "to be legislated rather than fought over."[55] We overcome skepticism by resisting the temptation to think that even "crazy" and quite untestable hypotheses might still be true because THE WORLD could just be that way:

> 'The perspective I urge with respect to all these cases
> is that there is nothing _hidden_ here, no noumenal fact
> of the entities' really being conscious or really not
> being conscious, or of the qualities' really being the
> same or really being different. There are only the ob-
> vious empirical facts: that rocks and nations are
> grossly dissimilar from people and animals;...Rocks and
> nations _aren't_ conscious; that is a fact about the notion
> of consciousness we actually have.   56

Endnotes - Chapter Three

1. Putnam, 1981a, pp. 49-50.

2. Ibid., p. ix.

3. Ibid., p. xii.

4. Ibid., p. x.

5. Ibid., p. xi.

6. Ibid., p. 41.

7. Ibid., p. 16.

8. Ibid., p. 5.

9. Ibid., p. 23.

10. Ibid., p. 10.

11. Ibid., p. 6.

12. Ibid., p. 8.

13. Ibid., p. 8.

14. Ibid., p. 14.

15. Ibid., p. 15.

16. Ibid., p. 15.

17. Ibid., p. 16.

18. Lewis, 1984; Smith, 1984.

19. Putnam, 1981a, pp. 32-33.

20. Ibid., p. 35.

21. Ibid., p. 33.

22. See Goodman, 1954, Chapter III.

23. Putnam, 1981a, pp. 46-48.

24. Ibid., p. 72.

25. Ibid., p. 73.

26. Ibid., pp. 73-74.

27. Ibid., pp. 53-54.

28. Ibid., p. 71.

29. Ibid., p. 90.

30. Ibid., p. 101.

31. Putnam, 1982b, p. 211.

32. Ibid., pp. 227-228.

33. Putnam, 1982d.

34. Ibid , p. 198.

35. Ibid., p. 200.

36. Putnam, 1982b, p. 273.

37. Loc. cit.

38. Ibid., p. 275.

39. Loc. cit.

40. Putnam, 1983a, pp. ix-x.

41. Ibid., p. xii.

42. Putnam, 1981a, p. 50.

43. Loc. cit.

44. Loc. cit.

45. Ibid., p. 52.

46. Ibid., p. 54.

47. Ibid., p. 55.

48. Ibid., pp. 55-56.

49. Field, 1982.

50. Putnam, 1981a, p. 61.

51. Ibid., p. 64.

52. Ibid., p. 73.

53. Ibid., p. 54.

54. Ibid., pp. 100-101.

55. Ibid., p. 92.

56. Ibid., p. 102.

Chapter Four   Putnam's Brain-Teaser

IV-a   The Problem Posed

1.   Metaphysical realism and skepticism cohabit in one of those tortured relationships beloved of modern dramatists - despite their apparent antagonism, they seem to need one another to retain a sense of their own significance. Confining our attention to one side of this unhappy drama, the realist has traditionally relied upon the skeptic to give substance to the idea that truth is, in Putnam's words, "radically non-epistemic," forever outstripping the reach of justification. What better model of truth so conceived, after all, than the skeptic's contention that even our firmest convictions might be mistaken in that we might be victims of demonic deception or of the machinations of some evil scientist? But the realist's right to this favourite model of truth, encapsulated in the claim that we might all be "brains in a vat", is challenged by Putnam's claim, in the opening chapter of RTH, that, given the realist notion of truth as correspondence to THE WORLD, and given certain plausible constraints on reference, we can know that we are not brains in a vat (or, more accurately, "brains in a vat" of a particular kind, as we shall see). If the conclusion of Putnam's argument is relatively clear, however, a brief survey of the secondary literature on this topic would indicate that the precise nature of the argument whereby Putnam purports to establish that conclusion, the validity of that argument, and the philosophical significance of the conclusion reached, are matters that have proved rather less transparent.

One writer who has attempted to dispel at least some of the gloom that threatens to enshroud these issues is Anthony Brueckner. Brueckner maintains[1] that, if we charitably overlook certain textual indications

to the contrary, Putnam's argument is one that admits of a sound formulation that appears to establish the desired conclusion. The argument, as construed by Brueckner, has the form of a dilemma, with the disjunctive premiss, "Either we are brains in a vat, or we are not brains in a vat." After a preliminary consideration of certain other critical responses to Putnam's argument, I shall argue, against Brueckner, that no such "dilemmatic" formulation of the argument constitutes a piece of sound reasoning that could be employed by a cognitive agent in the furtherance of achieving Putnam's conclusion. I shall further argue, however, that an alternative formulation of the argument will permit us to reason soundly to that conclusion. I shall conclude by reflecting briefly upon the intended philosophical significance of Putnam's discussion of brains in a vat.

2.   Let me begin by rehearsing certain general features of the argument in chapter 1 of RTH, features that I take to be relatively uncontroversial[2]. The premisses of the argument comprise a pair of "plausible" constraints on reference and a specification of what might be described as the "terms of envatment". First, the constraints on reference[3]:

(Ra) "Magical theories of reference" - theories that take reference to be an intrinsic relation between symbol and referent - "are wrong, wrong for mental representations and not only for physical ones."

(Rb) "One cannot refer to certain kinds of things, e.g. trees, if one has no causal interaction at all with them."

As to the terms of envatment, we are asked to entertain the possibility that the world is actually as portrayed in the following claims:

(Ta) All sentient beings are permanently isolated brains in a given vat.

(Tb) The automatic machinery controlling the vat is programmed so that the brains experience a hallucination qualitatively identical to our own experiences.

(Tc) The machinery is programmed to give us all a <u>collective</u> hallucination, rather than separate unrelated ones.

(Td) There neither is, nor ever was, an intelligent creator/designer of this state of affairs --rather, the universe just happens to be this way.

Following Brueckner, I shall employ the term 'BIV' as shorthand for the term 'brain in a vat in the circumstances outlined in (Ta)-(Td)'.

Putnam argues that the supposition that we are actually BIVs cannot possibly be true because it is "self-refuting". This claim seems counter-intuitive because there is surely a "physically possible world" in which all sentient beings are BIVs, and it would appear that if <u>we</u>, who (by Tb) share all experiences, "narrowly" construed, with the denizens of such a possible world, can give an argument to show that we are not BIVs, then so could <u>they</u>. Putnam responds that

> although the people in that possible world can think
> and "say" any words we can think and say, they cannot
> (I claim) <u>refer</u> to what we can refer to. In particular,
> they cannot think or say that they are brains in a vat
> (<u>even by thinking "we are brains in a vat"</u>).

For, given Ra and Rb, we cannot infer, from the fact that we share with the BIVs all mental states "narrowly conceived", that our verbal and visual representations refer to the same things. In fact, when a BIV <u>thinks</u>, "There is a tree in front of me", "his thought does not refer to actual trees... [but] to trees in the image, or to the electronic impulses that cause tree experiences, or to the features of the program

that are responsible for those electronic impulses."[5] Since this will

also apply to the uses of such words as 'vat' and 'nutrient fluid' in

"vat-English",

> it follows that if their "possible world" is really the
> actual one, and we really are the brains in a vat, then
> what we now mean by "we are brains in a vat" is that we
> are brains in a vat in the image or something of that
> kind (if we mean anything at all). But part of the hypo-
> thesis that we are brains in a vat is that we aren't
> brains in a vat in the image (i.e. what we are hallu-
> cinating isn't that we are brains in a vat). So, if we
> are brains in a vat, then the sentence "We are brains
> in a vat" says something false (if it says anything).
> In short, if we are brains in a vat, then "we are brains
> in a vat" is false. So it is (necessarily) false.   6

3.   At this point, philosophers tend to divide into two camps: those

who feel that there is obviously something wrong with Putnam's argument,

and those who feel that there is obviously something right with it.

Members of the former camp defend their intuitions on the matter by

attributing to Putnam an impressive array of fallacious or otherwise

inadequate lines of reasoning: members of the latter camp respond with

cries of misrepresentation. But such defenders of Putnam surely owe us

at least the outlines of an adequate and correct representation of the

argument. Can such a representation be given?

Central to Putnam's argument, obviously, is the claim that the

limited referential repertoire available to a BIV would prevent it from

representing to itself its own unfortunate situation. Let me briefly

comment on certain strategies that attempt to nip the argument in the

bud by rejecting this claim. Some writers (e.g. Alan Goldman[7] and David

Lewis[8]) have maintained that Putnam's claim rests upon a "strongly

causal theory of reference", and that we can avoid the conclusion of

the "brain in a vat" argument by the simple expedient of rejecting such a theory. (Goldman suggests that Putnam's use of a "strongly causal theory of reference" in the "brain in a vat" argument is itself a strong argument against such a theory!) One response to this sort of objection is to point out that Ra and Rb do not amount to a "strongly causal theory of reference". They merely impose certain necessary conditions on the referential roles played by certain terms, and these conditions, argued for in the opening pages of RTH, seem to be intuitively plausible constraints upon the adequacy of any theory of reference. Secondly, even if one countenances giving up Rb, this will not render Putnam's argument inconsequential to the extent that those philosophers who have employed "brain in a vat" hypotheses as models of externalist truth have themselves endorsed Rb.

A related externalist strategy is to accept the constraints on reference but propose some sort of causal mechanism whereby BIVs could obtain referential access to real objects. The problem with this strategy is that the proposed mechanisms tend to involve an implicit violation of the terms of envatment. For example[9], we might imagine that the brains' collective hallucination contains an episode in which a voice speaks from the sky and tells the "listeners" that they are really BIVs. It might be thought that, after suitable reflection on such an experience, the brains might be gradually initiated into the language of the outside world. But clearly, if reference to the outside world were to be established in such a case, it would not be merely as a result of such reflections on the BIVs' behalf - reflections that might equally well occur after vat-reading a vat-copy of RTH - but, rather, would depend upon the mediating activity of whatever intelligence was

responsible for the heavenly communication, assuming that the latter was _itself_ referentially connected to real vats and real brains. Thus any such proposal would violate the terms of envatment, which proscribe the mediating influence of external intelligence.

More interesting, perhaps, is Peter Smith's suggestion[10] that the realist's use of the "brain in a vat" hypothesis should not be taken as an attempt to provide a literal description of our existential situation, but as an attempt to characterise that situation by means of exemplification. The realist's claim, according to Smith, is that our _relation_ to THE WORLD might be the one exemplified _in the image_ by the relation between what we call hallucinated brains in a vat, on the one side, and the world in the image, on the other. The latter relation, which Smith terms "the delusive relation", may be the very relation in which we stand to THE WORLD. However, as Nelson Goodman has stressed[11], exemplification is a mode of symbolisation, and, as such, it presupposes the referential relation which Goodman terms "converse denotation". X can exemplify _f_ only if 'f', construed as a label, denotes X. In Smith's example, the relation "in the image" between a permanently hallucinated brain and the rest of the world can exemplify the delusive relation only if the former is denoted by the label, 'the delusive relation'. More significantly, the relation between ourselves and THE WORLD can exemplify the _same_ relation as this relation."in the image" only if 'the delusive relation' also denotes - that is, _refers_ to - _that_ relation. But why should an envatted brain's ability to use the label 'the delusive relation' to refer to a specific real-WORLD _relation_ between real brains and other things in THE WORLD be _any less problematic_ than its ability to use the word 'vat' to refer to real vats? Once we note that

exemplification presupposes reference, we will not be overly impressed by the suggestion that what a BIV cannot achieve by reference it can achieve by exemplification.

4. So much for the initial skirmishing. In what follows, I shall assume that Putnam is correct in claiming that, given the constraints on reference Ra and Rb, a BIV would be unable to represent to itself its own lot. The question, then, is whether, granting this claim and the two constraints on reference, there is an otherwise sound argument leading to the conclusion that we can know that we are not BIVs. There is general agreement among commentators that Putnam's reasoning rests upon the purportedly "self-refuting" nature of the supposition that we are BIVs. But, as commentators have also remarked, Putnam's explicit suggestions as to precisely how this supposition is "self-refuting" seem to generate arguments that are patently invalid. Jane McIntyre[12], for example, identifies and rejects two possible formulations of the argument suggested by certain of Putnam's remarks.

Firstly, we have the following formulation, based upon Putnam's contention that the "self-refuting" nature of the supposition that we are BIVs resides in the fact that "if we can consider whether it is true or false, then it is not true":

(1a) Since BIVs cannot refer to real vats and real brains, they cannot consider the hypothesis that they are real brains in real vats; so

(1b) If we can consider this hypothesis, then we are not BIVs; and

(1c) If we are not BIVs, then the hypothesis that we are BIVs is false.

The problem with this formulation of the argument, as McIntyre points out, is that we cannot detach the consequent of 1c unless we have the antecedent of 1b as an additional premiss. But, since a necessary con-

dition (by 1b) for the antecedent of 1b is that we are not BIVs, to admit such an additional premiss would be to beg the question at issue.

Alternatively, Putnam's claim that "if we are brains in a vat, then 'We are brains in a vat' is false" suggests that his argument has the following form: "p entails not-p; therefore not-p". More specifically:

(2a) If we <u>are</u> BIVs, then we can only refer to brains and vats "in the image" (or to electronic impulses or features of computer programmes), and thus our utterances of the the sentence, "We are BIVs", say something false;

(2b) If our utterances of the sentence, "We are BIVs", say something false, then we are not BIVs; so

(2c) If we are BIVs, then we are not BIVs (2a, 2b, H.S.); so

(2d) We are not BIVs.

McIntyre notes that the failing in this formulation of the argument is that it trades fallaciously on the bilinguality of the sentence, "We are BIVs". In so far as 2a is an extension of Putnam's claim that a BIV cannot describe its own lot, the antecedent of 2a hypothesises that we are <u>real</u> BIVs, and the consequent asserts the falsity of the vat-English sentence, "We are BIVs". If the inference of 2c from 2a and 2b is to be valid, then the antecedent of 2b must also assert the falsity of the vat-English sentence, "We are BIVs", and the truth of 2b will require that its consequent affirm that we are not BIVs "in the image". But, in that case, what 2c says, when properly expanded, is that if we are real BIVs, then we are not BIVs "in the image": and, since the consequent of 2c is then not the denial of its antecedent, the derivation of 2d will be invalid.

IV-b   Brueckner's "Dilemma"

Each of the above formulations incorporates elements from Putnam's own presentation of the "brain in a vat" argument, and each, taken by itself, is inadequate to establish the required conclusion. But perhaps we may hope to do better if we combine resources from each of McIntyre's versions into a single argument. In so doing, we may fasten upon another of Putnam's pronouncements concerning the "self-refuting" nature of the BIV hypothesis - his contention that "jt is <u>the supposition that the thesis is entertained or enunciated</u> that implies its falsity." Rather than construe this as an avowal of McIntyre's 1b, we may read it as the claim <u>that</u> any <u>utterance</u> or <u>supposition</u> of the sentence, "We are BIVs", is false, <u>whether or not we are BIVs</u>. The suggestion is that we treat the argument as a <u>dilemma</u>, with the disjunctive premiss, "Either we are BIVs or we are not BIVs", and with conditional premisses modelled after 2a and 1c. Let me refer to any version of Putnam's argument having this general structure as a "dilemmatic" formulation.

Perhaps the clearest presentation and defence of a dilemmatic formulation of the argument occurs in Anthony Brueckner's paper, "Brains in a Vat". Brueckner employs first-person singular throughout, where we have thus far employed first-person plural; but, given the generality condition built into our definition of a BIV, we may disregard this difference for present purposes. Brueckner offers the following formulation of the argument:

> (1) Either I am a BIV (speaking vat-English) or I am
>   a non-BIV (speaking English)
> (2) If I am a BIV (speaking vat-English), then my utter-
>   ances of "I am a BIV" are true iff I have sense-
>   impressions as of being a BIV.
> (3) If I am a BIV (speaking vat-English), then I do not

> have sense-impressions as of being a BIV.
> (4) If I am a BIV (speaking vat-English), then my utter-
> ances of "I am a BIV" are false. (2, 3)
> (5) If I am a non-BIV (speaking English), then my utter-
> ances of "I am a BIV" are true iff I am a BIV.
> (6) If I am a non-BIV (speaking English), then my utter-
> ances of "I am a BIV" are false. (5)
> (7) My utterances of "I am a BIV" are false. (1, 4, 6) 13

For convenience of exposition, I shall henceforth refer to members of
this set of sentences as B1, B2, etc.

If we grant, as I think we should, the second horn of the dilemma,
then the soundness of Brueckner's version of the "brain in a vat" argu-
ment turns on the sub-argument for the first horn, B2-B4. Brueckner
considers the following objection to B2. The sentence

> (T) My utterances of "I am a BIV" are true iff I am a BIV.

is true as uttered by a BIV in vat-English, just as it is true as
uttered by a non-BIV in English. Thus we seem to have

> (B8) If I am a BIV (speaking vat-English), then my utter-
> ances of "I am a BIV" are true iff I am a BIV.

But, Brueckner notes, B2 and B8 yield incompatible verdicts as to the
truth of "I am a BIV" in the case that I _am_ a BIV and B3 is true. His
response is to argue that we should reject B8 for the same sorts of
reasons that would lead us to reject

> (B9) If I am speaking a language in which 'tail' refers
> to legs, then horses have four tails.

He distinguishes two strategies that we might adopt in assessing the
truth-value of B9:

i/ We might "interpret its consequent in such a way that the condition
specified in its antecedent is taken as applying to the language used
in stating the consequent," or

ii/ We might determine whether, in a world in which the antecedent is

true, "the English sentence 'horses have four tails' would be true."
On the first strategy, B9 has a true consequent whenever it has a true
antecedent, whereas on the second strategy this is not the case.
Brueckner maintains that the second strategy is clearly the correct one,
and he then defends B2 on these grounds:

> The idea is that the consequent of [B2] gives the
> correct English specification of [the] truth-condit-
> ions [of "I am a BIV" as uttered in vat-English]. The
> consequent of [B8] if understood as a piece of vat-
> English, would give the correct specification of the
> mentioned sentence's vat-English truth-conditions. But
> the language used throughout [B8], we assume, is
> English.

In response to the further objection that this last assumption is
hardly legitimate in an argument that purports to establish that I am
a non-BIV speaking English, Brueckner maintains that the sentences
B1-B7 constitute a sound argument whether entertained by a BIV or by a
non-BIV. He argues for this conclusion by constructing an English rep-
resentation of the propositions that would be expressed by a BIV's
utterance of B1-B7: the resultant argument, he maintains, is a sound
one.

I have belaboured Brueckner's treatment of these issues because I
believe such matters are central to an assessment not only of Brueck-
ner's version of the "brain in a vat" argument, but also of "dilemmatic"
formulations in general. I shall argue that Brueckner's account of the
interpretation of indicative conditionals such as B2 is incorrect, and
that, once the correct interpretation is grasped, it is apparent that
no adequate dilemmatic formulation of Putnam's argument can be given.

Brueckner's argument for preserving B2 is broadly analogical in

nature. B8, which is incompatible with B2, is to be treated on a par
with B9. And B9, he further maintains, is no more plausible than its
counter-factual counterpart

> (B9*) If I were speaking a language in which 'tail' refers to
> legs, then horses would have four tails. .

To the extent that we endorse both the Kripke-Putnam account of refer-
ence and a possible-world semantics for counter-factuals, we will
surely be sympathetic to the claim that B9* is false. Should we, then,
pass similar judgment on its indicative counterpart B9?

To clarify the relationship between counter-factual and indica-
tive conditionals in these kinds of cases, it will be helpful to con-
sider a clearer example. I shall assume a familiarity with the general
outline of Putnam's "Twin Earth" examples. Consider the following pair
of conditionals:

> (TE) If I am on Twin Earth (speaking L_{TE}), then water is not
> H_2O.
> (TE*) If I were on Twin Earth (speaking L_{TE}), then water would
> not be H_2O.

Given the Kripke-Putnam account of reference and a possible-world
semantics for counter-factuals, TE* will be judged _false_ because:
(a) 'Water' designates in all possible worlds what it designates in
the actual world (by the account of reference for natural-kind terms);
(b) The actual world in which we are entertaining TE* is Earth (we
assume), and we are speaking L_E, not L_{TE}; (c) 'Water', in L_E, desig-
nates H_2O (ex hypothesi the "Twin Earth" example); so (d) water will
be H_2O in all possible worlds, and therefore, _a fortiori_, in all
possible worlds in which the antecedent of TE* is true.

If we try to use the same line of reasoning, _mutatis mutandis_,

to demonstrate the falsity of TE, however, we run into difficulties at step (b). For our use of the indicative rather than the counter-factual conditional conversationally implies an uncertainty as to whether the antecedent holds <u>in the actual world</u>, and thus an uncertainty as to which world <u>is</u> the actual world and which language we are speaking when we entertain the sentence TE. (Compare, for example, "If water is not H_2O, then water is not H_2O in any possible world" and its (false) counter-factual counterpart.) An individual entertaining the sentence TE must be one who, for whatever reason (choose your own favourite science-fiction device), is genuinely unsure as to the linguistic community to which she belongs - just as the person entertaining an indicative conditional whose antecedent is "I am a BIV" must be assumed to be genuinely uncertain as to her ontological status (prior to being rescued by Putnam, of course!).

Thus, in place of steps (b)-(d) in our evaluation of TE*, we have to consider the following possibilities:

(A) (b') The actual world is Twin Earth, and we are speaking L_{TE}; so

   (c') 'Water' in the actual world designates XYZ; so

   (d') Water is XYZ in all possible worlds, including the actual world, and the consequent of TE is true.

(B) (b") The actual world is Earth, and we are speaking L_E; so

   (c") 'Water' in the actual world designates H_2O; so

   (d") Water is H_2O in all possible worlds, and the consequent of TE is false in all possible worlds, including the actual world.

But since, in case B, the antecedent of TE is false in the actual world, it follows, given our analysis of case A, that TE is at least materially true. In fact, TE satisfies the stronger condition that, in

any world in which it is uttered or entertained, its consequent can be false in that world only if its antecedent is also false in that world.

Consider, on the other hand, the indicative counterpart of the true counter-factual

(TE≠*) If I were on Twin Earth (speaking L_{TE}), then water would be H_2O.

The sentence

(TE≠) If I am on Twin Earth (speaking L_{TE}), then water is H_2O.
need not express even a material truth. For if, as in scenario A above, the actual world in which I am entertaining TE≠ is Twin Earth, then water rigidly designates XYZ and TE≠ will have a true antecedent but a false consequent.

The conclusion I wish to draw from the preceding considerations is that the correct strategy to adopt in assessing the truth-value of indicative conditionals whose antecedents hypothesise that a particular language is being spoken is the first, rather than the second, of the two distinguished above. That is, we should interpret the consequent of the conditional as an utterance in the language hypothesised in the antecedent. If we apply this result to Brueckner's formulation of the "brain in a vat" argument. however, we must reject B2 and uphold B8, since it is the latter rather than the former that gives a correct formulation, in vat-English, of the vat-English truth-conditions for the sentence, "I am a BIV". But clearly, if we substitute B8 for B2 in Brueckner's argument, there is no route to B4, and we cannot establish the first horn of the dilemma. Thus Brueckner's version of the dilemmatic formulation fails. And there is no reason to think that this conclusion will not hold for any other version of the dilemmatic for-

mulation of Putnam's argument.

In fact, I think an even more radical conclusion may be warranted, namely, that, while there is a sense in which <u>B8</u> expresses a true proposition, a speaker cannot <u>express</u> that proposition by uttering the sentence B8. If this is correct, then I think the result generalises to any other indicative conditional sentences with antecedent "I am a BIV" that might be appealed to in an attempt to resuscitate the dilemmatic formulation.

It will be helpful, here, to introduce the following notational convention:

(a) Where a string of words is to be interpreted as an utterance in English, we may characterise it thus:- (_____)$_E$

(b) Where a string of words is to be interpreted as an utterance in vat-English (L_V), we may characterise it thus:- (_____)$_V$

(c) Where a string of words is to be interpreted as an utterance in what we may term "double-vat-English" (L_{VV}) - the language which speakers of L_V attribute to the entities denoted by <u>their</u> term 'BIV' - we may characterise it thus:- (_____)$_{VV}$ -

Now, we have seen that the consequent of an indicative conditional whose antecedent hypothesises that a given language is being spoken, is to be interpreted as an utterance in the language hypothesised when we assess the truth-value of the conditional. So, in the case of B8, it would seem that we must assess the consequent as an utterance in L_V. In worlds in which the antecedent of B8 is true, the consequent, so interpreted, will also be true. <u>But</u> in such worlds the language used in the <u>antecedent</u> of B8 is L_V, and to hypothesise in L_V that one is a BIV is to hypothesise that one is speaking L_{VV}. We may clarify what is

going on here by means of our notation. The following sentences express the true proposition that we intuitively associate with B8:

(P1) If (I am a BIV)$_E$, then (my utterances of 'I am a BIV' are true iff I am a BIV)$_V$

(P2) If (I am a BIV)$_V$, then (my utterances of 'I am a BIV' are true iff I am a BIV)$_{VV}$

But - and here's the rub - no speaker can use the sentence B8 to express either of the true propositions expressed by P1 and P2. For P1 and P2 are <u>bilingual sentences</u>. Only a person who could speak both English and vat-English, for example, could express what P1 expresses by uttering B8. But such bilingualism is obviously impossible, given the nature of the case. What a speaker <u>can</u> express by uttering B8 is either

(P3) If (I am a BIV)$_E$, then (my utterances of 'I am a BIV' are true iff I am a BIV)$_E$, or

(P4) If (I am a BIV)$_V$, then (my utterances of 'I am a BIV' are true iff I am a BIV)$_V$.

But both P3 and P4 are <u>false</u>. Taking P3, for example, in worlds in which I am a BIV it is not true that my utterances of "I am a BIV" are true iff I am a <u>real</u> brain in a <u>real</u> vat.

Note that this sort of difficulty does not arise in the sorts of Twin Earth examples discussed earlier, because, in such examples, there is no inconsistency in the supposition that the language <u>hypothesised</u> in the antecedent is also the language <u>spoken</u> in the antecedent. There <u>is</u> an inconsistency of this sort when the antecedent is "I am a BIV", however, and this is the root of our difficulties.

IV-c   The BIV Hypothesis as "Performatively Self-Refuting"

.   If I am right in claiming that the dilemmatic formulation of the
"brain in a vat" argument is fundamentally flawed, then, given the
opinion current in the literature that this is the only alternative
to the patently inadequate versions of the argument canvassed by
McIntyre, we may feel that Putnam's critics have been right all along
on this issue. But I think such a conclusion would be premature. Our
discussion thus far suggests that an adequate version of Putnam's
argument must satisfy at least the following requirements:

(1) It must not employ "bilingual" premisses, nor must its soundness
in any way presuppose bilinguality, and

(2) It must not depend for its soundness on any premisses which would
be available to us only if we are not BIVs.

   But, if the argument is to be both valid and restricted, in its
premisses, to the use of true sentences available to both BIVs and non-
BIVs, then it seems that, if the argument is to be sound, it must be
sound when interpreted in both English and vat-English. And this raises,
yet again, the objection that, if BIVs themselves can give the argument
that proves one is a non-BIV, this demonstrates that there must be some-
thing wrong with the argument. I think that to see why this objection
is misconceived is to see how the argument should be formulated. If
the argument is to be sound in both languages, then this suggests that
it must turn upon some feature which the sentence, "I am a BIV",
possesses independently of the linguistic context of utterance. This,
I believe, is the correct way to read Putnam's sibylline pronouncement
that the "brain in a vat" hypothesis is "self-refuting" in the sense

that "it is the supposition that the thesis is entertained or enunciated that implies its falsity." The crucial feature of the sentence, "I am a BIV", I suggest, is that it is "performatively self-refuting" -that is, it is false whenever it is asserted in virtue of preconditions for its being asserted[16]. "I am not here now" is a classic example of a performatively self-refuting sentence. It contains three terms whose precise reference will vary according to circumstances of utterance, but we do not need to know the context of assertion to know that a speaker who utters the sentence (under "standard conditions" - see below) is making a false assertion. If it could be shown that the sentence, "I am a BIV", is similarly performatively self-refuting, then we would not need to know the linguistic context of utterance to know that a speaker who utters the sentence is stating a falsehood. And, in so far as the speaker in question is oneself, one could know that one was not a BIV.

Can we, then, establish that "I am a BIV" is performatively self-refuting without falling foul of any of our earlier difficulties? I think something along the following lines will do the trick:

(3a) The sentence, "I am a BIV", can be used to make a true assertion only by one who satisfies both of the following conditions:

    (i) s/he is a brain in a vat;

    (ii) s/he is able to use the terms 'brain', 'vat', etc., to describe her/his own situation.

(3b) Given constraints Ra and Rb, anyone who satisfies 3a/i cannot also satisfy 3a/ii - so no-one could satisfy both conditions.

(3c) Therefore no-one could use the sentence, "I am a BIV", to make a true assertion because of its preconditions of utterance, and

the sentence is therefore performatively self-refuting. So

(3d) I cannot be making a true assertion when I utter the sentence,

"I am a BIV", whatever language I am speaking. So

(3e) I know I am not a BIV.

Let me briefly respond to three possible sources of unease with
this formulation of the argument. Firstly, unpacking the argument in
3b obviously requires that we have available to us a distinction between
BIVs and non-BIVs, English and vat-English, etc. It might be thought
this violates one or other of the two requirements noted at the begin-
ning of this section, either by requiring that we be bilingual, or by
requiring that we be able to speak "real" English and be able to refer
to "real" brains. But the requisite distinction between real brains
and real vats, on the one hand, and brains in the image and vats in the
image, on the other, is available to speakers of both English _and_ vat-
English - it _must_ be, because the two languages are _lexically_ and _synt-
actically_ identical. The difference between them is semantic - which
is to say that, for speakers of vat-English, the referents of the terms
'real vat' and 'vat in the image' _will both be things in the image_;
while, for speakers of English, the referents of the same two terms
will both be things in the real world. Thus our ability to avail our-
selves of the distinctions required to spell out the argument in 3a
and 3b in no way prejudges the question whether we are understanding
that argument in English or in vat-English.

Secondly, it might be claimed that 3a is _false_ as it stands, for
it imposes no restrictions on the admissible languages to which the
sentence, "I am a BIV", may belong. There is obviously a need to exer-
cise considerable caution when making the claim that a particular sen-

tence is performatively self-refuting. On the one hand, one must rel-
ativise one's claim to "standard conditions of utterance" for the sen-
tence in question in order to rule out counter-examples like the
following: I utter the sentence, "I am not here now", but I change my
position in the room in between uttering the word 'here' and uttering
the word 'now'[17]. On the other hand, one must rule out utterances of
the sentence that are tokens in languages having a radically different
semantics from the language to which one takes the sentence to belong:
for example, one must rule out languages in which the string, "I am
not here now" has the same interpretation as the English sentence,
"I am bald". In the present case, we might consider a possible world
in which the language spoken is identical to English save that, where
we use the terms 'brain' and 'vat', they use the terms 'professor'
and 'university'. and vice versa. If premiss 3a is permitted to encom-
pass the denizens of such a possible world, then it would seem to fail
to give a correct characterisation of the conditions under which the
sentence, "I am a BIV", can be used to make a true assertion.

Normally, it is relatively easy to avoid this kind of problem -
one need only adopt the simple expedient of relativising one's claim
that a given sentence is performatively self-refuting to utterances
of that sentence in a given language. But, in the present case, it
is essential that our claim encompass utterances of the sentence in
question in more than one language. We could relativise the claim to
the two languages English and vat-English, but this manouevre looks
suspiciously ad hoc. Fortunately, however, there is an easier solution
to the problem for the restriction we seek is already implicitly con-
tained in the terms of envatment used to define a 'BIV'. Condition Tb

stipulates that the brains in a vat experience a hallucination qual-
itatively identical to our own experience - in Putnam's terminology,
the brains have the same "notional world" as we do. One feature of our
notional world is the experience of tokening sentences in response to
other experiences: for example, we token the sentence, "I see a vat
before me", in response to vat-before-me experiences. But the beings
in our possible world who use the word 'brain' where we use the word
'professor' will not share this feature of our notional world. In their
notional world, the sentence "I see a vat before me" is tokened in
response to a professor-before-me experience.

Thirdly, we need to defend the inference from 3d to 3e against
Brueckner's contention that the most that Putnam can establish is 3d.
I have already argued against Brueckner's claim that 3d can be estab-
lished by a dilemmatic form of reasoning, but his argument against the
inference from 3d to 3e is quite independent of this claim and there-
fore requires separate consideration. He argues that, to the extent
that we remain ignorant as to whether we are actually speaking English
or vat-English, we do not know what proposition we are asserting when
we utter the sentence, "I am a BIV", even though we know (by 3d) that
the proposition asserted is false. He claims that the situation that
we are in at step 3d is analogous to that of an individual who is ig-
norant of set theory but who is informed, by one known to be an author-
ity on such matters, that "Omega is not a regular cardinal". Such an
individual, Brueckner maintains, may properly be said to know that the
sentence, "Omega is not a regular cardinal", expresses a true proposi-
tion, but it would be wrong to credit the individual with knowledge
that omega is not a regular cardinal because the meaning of the sentence

is not understood. Similarly, our ignorance as to which language we are speaking makes it incorrect for us to claim knowledge that we are not brains in a vat, because we do not understand what proposition the sentence, "I am a BIV", expresses. Thus we cannot infer 3e from 3d.[18]

We may note, firstly, that the difficulty urged upon us by Brueckner will infect all claims to knowledge that require, in their formulation, the use of terms whose interpretation in vat-English differs from their interpretation in English - which is to say, all claims to knowledge that go beyond the reporting of our mental states "narrowly" construed. This by itself may give us pause to reflect. And a little reflection will, I think, reveal the flaw in Brueckner's analogical argument. For the reason why we resist the inference from the individual S's knowledge that the sentence, "Omega is not a regular cardinal", expresses a truth, to S's knowledge that omega is not a regular cardinal, is surely that S does not understand the terms employed in the sentence in question: and this failure of understanding manifests itself in S's inability to employ those terms correctly in other contexts and to recognise when they do and do not apply. But there is no analogous failure of understanding in the case of the terms employed in the sentence, "I am a BIV". The relevant terms ('brain', 'vat', etc.) are ones that we are able to use correctly in all manner of different contexts to pick out those entities that are the referents of those terms in our language. Our ignorance consists in our not knowing the ultimate ontological status of the entities that we correctly identify as the referents of the terms in our language. But not knowing this is not sufficient reason to deny us understanding of that language,

given our transparent ability to use that language correctly. And, if
we do understand our language, then there is no reason to deny the
inference from our knowing that an understood sentence of that language
is true to our knowing the proposition expressed by that sentence. Thus
the inference from 3d to 3e is quite legitimate.

IV-d What the BIV Argument Doesn't Show

I began by remarking upon the unholy alliance between metaphysical realism and skepticism. It is to this alliance that I shall return in some concluding observations concerning the intended philosophical significance of the BIV argument. Commentators have tended to assume that Putnam's intention is to assail one or other of the partners in this alliance. Brueckner himself, among others[19], assumes that the target is the skeptic. But, as he himself points out, if this _were_ the object of the exercise, the skeptic has an easy answer to Putnam. For, as was apparent in our earlier consideration of the claim that a BIV would be incapable of describing its own lot, the argument requires the extreme isolation stipulated in the terms of envatment if we are to rule out possible causal channels between an envatted brain and the outside world. But the skeptic has nothing invested in so extreme a scenario. His purposes can be equally well served by a mode of envatment considerably more relaxed. And, if we relax the terms of envatment, no longer have the resources to exclude the possibility of reference to the outside world, and the consequent ability of an envatted brain to truly characterise its own situation. Charity would suggest that, in the absence of textual evidence to the contrary, we refrain from reading chapter 1 of RTH as an assault upon skepticism. In fact, "textual evidence" clearly indicates that such a reading would be incorrect: see, for example, page 6 of RTH, where Putnam explicitly distances his concerns from the issue of skepticism.

For analogous reasons, I think we should resist the suggestion by David Lewis[20] that, Putnam's more modest claims notwithstanding, the opening chapter of RTH is to be read as an argument against metaphys-

ical realism, serving as a companion to the "model-theoretic" argu-
ment. Putnam's intention, according to Lewis, is to generate a "causal-
theoretic" argument against the possibility of radical deception by
extending the line of reasoning employed with respect to the BIV hypo-
thesis to all of the other beliefs of an envatted brain. The idea is
that "the causal theory of reference can be used over and over...to
exonerate the brain from all accusations of error whatsoever (except
when the brain falls short of 'epistemic truth')."[21] Lewis argues that
Putnam is mistaken in thinking that the "brain in a vat" reasoning can
be so extended because "the requisite causal theory of reference does
not exist." I think Lewis' argument fails, partly because, as noted
above, Putnam's reasoning doesn't rest upon a "strongly causal theory
of reference", but I shall restrict my attention, in the present con-
text, to Lewis' construal of Putnam's intentions. Could what Lewis
terms the "causal-theoretic" argument serve alongside the "model-theo-
retic" argument in the battle against the externalist view of truth,
and does Putnam intend the discussion of the BIV hypothesis to serve
in such a role?

Prima facie, it would seem that the "brain in a vat" argument,
whether or not "extended", is at best neutral on the question whether
we could be radically deceived under "epistemically ideal conditions".
For, we may recall, Putnam's conclusion is not merely that a perman-
ently envatted brain could not be mistaken in certain (or all) of its
beliefs, but also that we can know that we ourselves are not BIVs. So
the fact, if fact it be, that a BIV could not have any beliefs that
were "epistemically true" yet really false appears to offer no grounds
for thinking that we might not be radically deceived, even under "epis-

temically ideal conditions". But it might be thought that this is to miss the broader significance of the "extended" "brain in a vat" argument, for the latter can serve as a model for anti-realism just as the BIV hypothesis has traditionally served as a model for realist truth. The anti-realist can maintain that our relationship to THE WORLD, on an externalist account, is analogous to the relationship between a BIV and the real world that contains it - in neither case is the relationship sufficient to establish a referential relation between the two parties. Thus states of THE WORLD cannot falsify those of our beliefs that are "epistemically true", any more than states of affairs in the real world can falsify the "epistemically true" beliefs of a BIV.

However, even if this analogy be found instructive or illuminating, it in no way constitutes an _argument_ against realist truth. For nothing in the "brain in a vat" argument, "extended" or not, requires that we see our epistemic situation as analogous to that of a BIV. As we have noted, the "epistemically true" beliefs of an envatted brain enjoy immunity from error only in so far as that brain is subject to the extreme isolation stipulated in the terms of envatment, which effectively excludes the possibility of a suitable causal link between particular terms in vat-English and particular items in the real world. But the causal constraints on reference, Ra and Rb, permit us to hypothesise an envatted brain _not_ subject to such terms of envatment, and able, as a consequence, to refer to items in the real world and to hold beliefs that are "epistemically true" yet _false_ in virtue of states of that world. It is _this_ hypothesis, the externalist will claim, that correctly models our epistemic situation, rather than the BIV hypothesis entertained by Putnam. We should hold, in other words, that terms in our

language stand in appropriate causal relationships to entities in THE WORLD, and that, as a consequence, our "epistemically true" beliefs may be false in virtue of states of THE WORLD. To counter this sort of externalist response, Putnam must question the implicit appeal to causal structure in THE WORLD itself, or to our understanding of a causal theory of reference capable of describing relations between language and THE WORLD. But these challenges to externalism require the richer resources of the "model-theoretic" argument: they cannot be funded by the causal constraints on reference and the "brain in a vat" argument. The latter, by themselves, give us no reason to prefer Putnam's BIV hypothesis to the externalist's counter-hypothesis of an envatted brain causally connected in an appropriate fashion to the real world, as a model of our own epistemic situation. Thus the extended "brain in a vat" argument cannot play the anti-externalist role ascribed to it by Lewis.

Nor, I think, is there reason to think that Putnam intends the discussion of the BIV hypothesis to play such a role. His explicit claims in this regard are relatively modest. At the beginning of chapter 3 of RTH, where the distinction between the externalist and internalist perspectives is first introduced, he asserts that "the question of 'Brains in a Vat' would not be of interest, except as a sort of logical paradox, if it were not for the sharp way in which it brings out the difference between these philosophical perspectives."[22] The internalist, who rejects the "God': Eye" point of view and holds that truth consists in some form of rational acceptability from "the various points of view of actual persons, reflecting various interests and purposes that their descriptions and theories subserve," can dismiss the BIV hypothesis on the grounds that it does not even represent a possible world: "The

supposition that there could be a world in which all sentient beings are Brains in a Vat presupposes from the outset a God's Eye view of truth, or, more precisely, a No Eye view of truth - truth as independent of observers altogether."[23] The externalist, on the other hand, believes that truth is independent of observers altogether, in which case "the hypothesis that we are Brains in a Vat cannot be dismissed so simply." The externalist seems committed to the view that the BIV hypothesis might conceivably be true, yet Putnam's argument purports to prove that the hypothesis cannot be true if reference is taken to be subject to certain causal constraints.

But to say that the externalist cannot dismiss the BIV hypothesis "so simply" is not to say that he cannot accomodate Putnam's argument without immediate prejudice to his position. The import of the "brain in a vat" argument for externalism is, I think, of a more subtle nature. In chapter 4 of RTH, Putnam rehearses a number of other skeptical challenges that the externalist must take seriously, as representing possible states of THE-WORLD - for example, the possibility that rocks are conscious. The internalist, Putnam maintains, can dismiss such "possibilities" on the grounds that they are rationally unacceptable given the concepts - our concepts - in terms of which they are formulated. The content of our concept of consciousness, for example, is to be cashed out in terms of the ways in which we use our language, not in terms of some "hidden" fact about which things are "really" similar to other things - "really" similar in virtue of the nature of THE WORLD. In a sense, the "brain in a vat" argument prefigures this line of reasoning, not only by countering a skeptical challenge by appealing to the limitations imposed by semantic considerations, but also by pointing,

paradoxically, to the intimacy between externalism and skepticism while appearing to sunder that intimacy.

Let me close by noting one further respect in which the opening chapter of RTH prefigures the larger themes of the book. One of those themes is an opposition to what Putnam terms "scientism", and a plea for a Goodmanian pluralism. The "brain in a vat" argument contributes to this larger theme by illustrating the limitations of science as an arbiter of truth and knowledge:

> There is a "physically possible world" in which we are
> brains in a vat ; what does this mean except that there is
> a description of such a state of affairs which is compatible
> with the laws of physics? Just as there is a tendency
> in our culture to take physics as our metaphysics, that
> is, to view the exact sciences as the long-sought descrip-
> tion of the "true and ultimate furniture of the universe",
> so there is, as an immediate consequence, a tendency to
> take "physical possibility" as the very touchstone of
> what might really actually be the case. Truth is phys-
> ical truth;   possibility physical possibility; and
> necessity physical necessity, on such a view. But we
> have just seen, if only in the case of a very contrived
> example so far, that this view is wrong. The existence
> of a "physically possible world" in which we are brains
> in a vat (and always were and always will be) does not
> mean that we might really, actually, possibly be brains
> in a vat. What rules out this possibility is not phys-
> ics but philosophy.

Endnotes - Chapter Four

1. Brueckner, 1986.

2. Although, as we shall see, even these features are misconstrued by certain critics. The most flagrant example, perhaps, is Jonathan·Harrison's 1985, where the argument rests not only upon the violation of the "terms of envatment" - see below - but also upon taking the problematic assertion to be a third-person claim ("S is a brain in a vat") rather than a first-person claim (": am a brain in a vat")

3. Putnam, 1981a, p. 16.

4. Ibid., p. 8.

5. Ibid., p. 14.

6. Ibid., p. 16.

7. Alan H. Goldman, 1984.

8. Lewis, 1984.

9. I owe this example to John Collier.

10. Smith, 1984.)

11. Goodman, 1968.

12. McIntyre, 1984.

13. Brueckner, 1986, p. 154.

14. Ibid., p. 158.

15. See Putnam, 1975c.

16. I think this is the reading of the argument that we get if we take the characterisations cited earlier in the context of the example of a "self-refuting" proposition offered by Putnam. Consider this passage, from pages 7-8 of Putnam, 1981a:
A "self-refuting supposition" is one whose truth implies its own falsity. For example, consider the thesis that all general statements are false. This is a general statement. So if it is true, then it must be false. Hence, it is false. Sometimes a thesis is called "self-refuting" if it is the supposition that the thesis is entertained or enunciated that implies its falsity. For example, "I do not exist" is self-refuting if thought by me (for any "me")... What I shall show is that the supposition that we are brains in a vat has just this property.... '

17. See Kaplan, 1977, for a discussion of these points.

18. Brueckner, 1986, pp. 164-167.

19. See, for example, Feldman, 1984, and Stephens and Russow, 1985.

20. Lewis, 1984..

21. Ibid., p. 235.

22. Putnam, 1981a, p. 49.

23. Ibid., p. 50.

24. Ibid., p. 15.

Chapter Five The "Model-Theoretic" Argument

V-a· The Argument in Context

1) "Sophisticated Realism" Revisited

As the exegetical explorations of Part I revealed, the principal
weapon in Putnam's arsenal in his assault upon externalism has been
the so-called "model-theoretic" argument, albeit under a number of
different guises. In this chapter, I shall examine and critically assess
various strategies which have been employed in attempts to counter this
argument. Since I wish to argue that much of the critical response to
the "model-theoretic" argument has been misdirected because of a failure
to attend with sufficient care to the dialectical context of the argu-
ment, it will be helpful to begin by recalling certain salient features
of that context. More specifically, I shall rehearse, in outline, the
nature of the "sophisticated realism" defended in the Locke Lectures and
"Reference and Understanding" and its bearing upon the distinction bet-
ween internal and metaphysical realism.

"Sophisticated realism", as expounded in Putnam's "pre-critical"
writings, is "an over-arching empirical hypothesis". The realist hypo-
thesis is that terms in mature sciences refer to real entities in the
world, and that the laws of theories in such sciences are typically
approximately true. More generally, the realist claims that there is a
correspondence relation between the words in a language and things in
the world, and between the sentences of that language and states of
affairs. "Truth", for the "sophisticated realist", is not to be identi-
fied with warranted assertibility: rather, it is recognised that our
beliefs might be warrantedly assertible yet false. The realist maintains
that realism, as an empirical hypothesis of this sort, is both implied

by our scientific picture of the world and justified by its capacity to explain our success as language-users, both in science and in our every-day commerce with the world. In the first place, our scientific picture includes an account of how knowledge, including scientific knowledge, is obtained. Knowledge, on this account, is acquired through causal inter-action with the world. The methods of the sciences are taken to be a reliable, but by no means infallible, means of arriving at knowledge of the world. It is because of the fallibility of even our best methods of inquiry - a fallibility recognised in the theories developed by the very disciplines that employ those methods - that the "warranted assertibil-ity" of a scientific hypothesis is compatible with its being false. Secondly, the realist's "over-arching empirical hypothesis" explains both the general contribution of language-use to the achievement of our goals and the fact that mature sciences exhibit "convergence". As to the latter, realism explains why the practice of scientists who proceed on the assumption that realism is correct should be attended by success, measured in terms of increasing empirical adequacy and instrumental efficacy of theory.

The sophistication of the "sophisticated realist" also extends to the recognition - again as something which follows from our general scientific picture - that there may be more than one true theory of the world. The realist sees our situation as analogous with that of the cartographer who accepts, as something which follows from her descrip-tion of the earth, the possibility of alternative correct mappings. Theories that are intuitively different in that they seem to ascribe incompatible properties to the world may, nonetheless, be equally true to the extent that they provide "equivalent descriptions" of the world.

Two theories may be regarded as "equivalent descriptions" if (a) they agree about certain features of the world, and (b) it follows from our comprehensive theory of the world that these are the only respects in which such theories should agree. The features which equivalent descriptions have in common may be termed "invariants". It is the mark of the "sophisticated realist" that, recognising the possibility of equivalent descriptions, he adopts a realist stance only with respect to those features which are invariants for the theories in a given scientific domain. The residual differences between theories that are equivalent descriptions - the respects in which such theories intuitively differ from one another - are properly understood not as ascriptions of different characteristics to the world, but, rather, as different ways of describing or explaining the same characteristics - that is, the invariants upon which the theories agree.

As we saw in chapter two above, Putnam's presentation of the distinction between "internal" and "metaphysical" realism in "Realism and Reason" is initially puzzling because he identifies internal realism with the "sophisticated realism" of the "pre-critical" writings (now termed "empirical realism"), whereas the latter position is clearly a form of metaphysical realism. We noted that our puzzlement stems from the fact that the central terms in "sophisticated realism" - terms such as 'reference', 'correspondence', and 'the world' - admit of both an internal and a metaphysical realist interpretation. Metaphysical realism purports to provide a model of what it is for any particular theory to be correct. The metaphysical realist assumes that there is a determinate and unique relation of reference between terms in a language and elements,

or sets of elements, in THE WORLD, where the latter comprises some fixed totality of mind-independent entities. A "correct" theory is one which is "true", where truth consists in a correspondence relation, mediated by reference so construed, between sentences in a language and states of THE WORLD. THE WORLD's independence of our representational capacities confers upon metaphysical realist truth the property of being "radically non-epistemic" in the sense that even an "Ideal Theory" might be false. Insofar as this model is to serve for all theories, the notions employed in characterising the model cannot receive their interpretation from within any particular theory, but must somehow be understood independently. In the "pre-critical" writings, Putnam argues for "sophisticated realism" as a correct theory in the metaphysical realist sense: speakers construct representations of THE WORLD, and these representations refer to and are approximately true of THE WORLD. Internal realism, on the other hand, "employs a similar picture within a theory". The world, of which speakers construct representations, is the world as we understand it through our various representations, and it is to the objects described within these representations that our terms refer. The "causal interactions" with the world through which knowledge is acquired are themselves "part of the subject-matter of the representation" that is our scientific account of knowledge.

2) Auxiliary Arguments

Before turning to the "model-theoretic" argument, I shall briefly consider two auxiliary arguments against metaphysical realism which Putnam offers in "Realism and Reason". The first, and most developed, argument addresses the "pre-critical" contention that the "sophisticated realist" can evade difficulties which beset his "naive" confrere by recourse to the doctrine of "invariants" and "equivalent descriptions". Putnam hypothesises a simple universe where THE WORLD is a straight line. We have two stories about THE WORLD that agree on the existence of "line segments", but disagree as to whether there are also "points", construed as infinitely small parts that stand to line segments in the same relation that the latter stand in to the line as a whole (or to bigger line segments). Story 1 affirms the existence of points, while Story 2 maintains that points are simply "logical constructs out of line segments". Putnam offers the following summary of his "pre-critical" analysis of such a situation:

> A "hard-core" realist might claim that there is a "fact of the matter" as to which is true - Story 1 or Story 2. But "sophisticated realists", as I have called them, concede that Story 1 and Story 2 are "equivalent descriptions". In effect, this concedes that line segments are a suitable set of "Invariants" - a description of THE WORLD which says what is going on in every line segment is a <u>complete</u> description. In the past, I argued that this is no problem for the realist - it's just like the fact that the earth can be mapped by different projections, I said....
> In particular, I believed, it can happen that what we picture as "incompatible" terms can be mapped onto the same real object - though not, of course, within the same theory...It is a property of the world itself, I claimed - i.e. a property of THE WORLD itself - that it "admits of these different mappings".   1

He now maintains, however, that this strategy allows the realist

to retain THE WORLD only at "the price of giving up any intelligible notion of how. THE WORLD is." The problem is that "any sentence that changes truth-value upon passing from one correct theory to another correct theory - e.g. an equivalent description - will express only a theory-relative property of THE WORLD." The more "Stories" we can come up with that can be regarded as "equivalent descriptions", the fewer the invariants that all correct theories will share, and the more properties of THE WORLD that will turn out to be theory-relative. Putnam further maintains that this analysis extends to actual physical theory:

> One can construe space-time points as objects, for ex-
> ample, or as properties. One can construe fields as
> objects, or do everything with particles acting at a
> distance (using retarded potentials). The fact is, so
> many properties of THE WORLD - starting with just the
> categorial ones, such as cardinality, particulars, or
> universals, etc. - turn out to be "theory-relative"
> that THE WORLD ends up as a Kantian "noumenal" world,
> a mere "thing-in-itself". If one cannot say how THE
> WORLD is theory-independently, then talk of all these
> theories as descriptions of "the world" is empty.

Let me make two brief comments on this argument. Firstly, as Putnam recognises, this is essentially Nelson Goodman's long-standing objection to the standard correspondence theory of truth[3]. Goodman draws the conclusion that the notion of THE WORLD, and of truth as correspondence to THE WORLD, have no useful role to play in philoso- phical reflection upon our cognitive practices and their fruits. It is not clear, however, that the argument supports Putnam's conclusion that the metaphysical realist picture is "incoherent"[4]. Secondly, the argument goes through only if we reject the contention of the "sophis- ticated realist" that there are constraints on the proper set of "invar-

iants" transcending our capacity to produce "Stories" that agree or disagree with one another in certain respects. For the "sophisticated realist", the _true_ set of invariants is given _not_ by the intersection, if such there be, of all the "rationally acceptable" stories that we treat as equivalent descriptions, but, rather, by the (set of) stories that capture how THE WORLD actually is. Thus the realist may acknowledge, in Putnam's argument, a correct portrayal of our _epistemological_ situation (or, perhaps, of our _present_ epistemological situation), yet maintain that there is an independent "fact of the matter" as to what the _true_ invariants are, and thus as to which "rationally acceptable" stories are in fact equivalent descriptions of THE WORLD.

The second of Putnam's arguments draws upon Quine's thesis of "ontological relativity"[5]. Quine argues that, when we engage in the enterprise of translating one language into another, the extension to be assigned to a given term of divided reference in the "foreign" language is determinate only relative to a chosen translation manual. As a result, "there are often _inequivalent_ relative interpretations of one theory in another." But, if we accept the metaphysical realist picture, then, Putnam maintains, there must be a correct mapping of the language of each theory into THE WORLD, and, if the theories are com-plete, "a _unique_ reference-preserving "translation" connecting the languages." Thus the metaphysical realist must hold, in the face of onto-logical relativity, that there is a fact of the matter as to "_which_ translation _really_ preserves reference in every such case!" The realist might argue that "the language has _more than one_ correct way of being mapped onto THE WORLD (it must, since it has more than one way of being correctly mapped onto a language which is _itself_ correctly mapped onto

THE WORLD)." But the metaphysical realist picture cannot survive such
a concession, Putnam argues, since we would now have the consequence
that "what is a unique set of things within a correct theory may not
be a unique set of things 'in reality'."[6] Again, however, this line
of reasoning seems open to the response that there is an illicit in-
ference from facts about our epistemological situation to metaphysical
conclusions. The "sophisticated realist" will presumably resist the
elision of the distinction between (i) the sense of "correctness"
according to which it is true that there is more than one "correct"
mapping of L_1 onto L_2, and (ii) the sense of "correctness" according to
which L_1 is itself "correctly" mapped onto THE WORLD. The realist will
contend that in the second of these senses - the sense in which "corr-
ectness" is a matter of corresponding to the way THE WORLD is - there
is only one correct way of mapping L_1 onto L_2, even if we don't know
which mapping is the correct one.

3) Interpreting the "Model-Theoretic" Argument

The "model-theoretic" argument challenges the metaphysical realist (externalist) thesis that truth is "radically non-epistemic". It is this thesis - the claim that even an "Ideal Theory", one that satisfied the highest possible standards of rational assertibility, might·still be false - that distinguishes externalist truth from Peircean truth. Putnam argues that the externalist is ·not entitled to assume the existence of something that is a precondition for the possible falsity of an "Ideal Theory", namely, a unique and determinate relation of·reference between terms, in a language and entities in THE WORLD that allows states of THE WORLD to confer determinate truth-values on sentences containing such terms, and thereby to confer the determinate truth-value "false" upon some sentence belonging to the "Ideal Theory". As we saw in chapter two above, the argument, formulated from within the externalist framework, runs roughly as follows. Assume the externalist conception of THE WORLD and of truth as reference-mediated correspondence to THE WORLD. Then, given an "Ideal Theory" of an appropriate cardinality:

(a) There is a mapping, and hence a "reference" relation, from the terms of the language of the theory, IT, onto elements, or sets of elements, in THE WORLD, such that THE WORLD satisfies IT - call this mapping SAT; and

(b) Nothing in our use of the language can serve to exclude this mapping as an "unintended" interpretation, since IT satisfies, ex hypothesi, all operational and theoretical constraints, and since SAT also satisfies our general intention to refer in such a way that our beliefs (for the most part) come out true.

(c). Furthermore, there are no other constraints available to the externalist that can rule out SAT as an "unintended" interpretation.

(d) So, if truth consists in a reference-mediated correspondence to THE WORLD, then an "Ideal Theory" cannot fail to be true.

.The foregoing summarises the argument presented in "Realism and Reason", an argument I set out in more detail in chapter two above. Before examining the critical response to the "model-theoretic" argument, however, it is important that we clarify certain details, and remedy certain common misconstruals, of the argument. It will be helpful, in this context, to draw upon the fuller presentation of the argument in "Models and Reality" which was also sketched in chapter two. The thrust of the argument of the latter paper, we may recall, is that, given a set of operational constraints formulated in terms of one vocabulary - what is, relatively speaking, the "observational" vocabulary - we cannot single out a unique, "interpretation-independent" reference for the terms in another vocabulary - relatively speaking, the "theoretical vocabulary" - by means of such operational constraints in conjunction with a set of theoretical constraints.

One matter upon which "Models and Reality" provides needed clarification is the sense to be given to the requirement, on an "Ideal Theory", that it satisfy "the set of all operational constraints". Some commentators, perhaps misled by the relative silence of "Realism and Reason" on this point, have assumed that a theory satisfies all "operational constraints", in Putnam's sense, if it is consistent with all "observational" evidence that will ever be available in practice - that is, all such evidence that we will ever in practice be able to accumulate by dint of our investigative endeavours.[7] Such a reading

threatens to launch a veritable fleet of counter-examples. Fortunately, we may have recourse, at this point, to a clearer exposition of what Putnam has in mind in "Models and Reality". In the passage in question, he is characterising the set of operational constraints formulated in terms of an "observational vocabulary" that refers to ordinary "middle-sized" objects, their properties, and events in which they partake. The set of operational constraints is to be constructed on the basis of the following: (a) an "observational vocabulary" containing terms such as 'red', 'touches', 'hard', and 'push'; (b) a set S - assumed to exist whether we can define it or not - comprising all things and events observable with the human sensorium; and (c) "a valuation...which assigns the correct truth-value to each n-place 0-term (for n = 1, 2, 3,....) on each n-tuple of elements of S on which it is defined."[8] Thus the set of operational constraints is to incorporate <u>all observable occurrences</u>, regardless of whether the necessary observations are ever actually made in practice.

"Models and Reality" also provides a helpful amplification of Putnam's somewhat peremptory dismissal, in "Realism and Reason", of the suggestion that a "causal theory" of reference might furnish the additional constraints on reference that the externalist requires. His claim, in the latter paper, is that the problem of determinate reference arises as much for the term 'causes' as it does for any other term. In the former paper, the point is made as follows:

> The problem is that adding to our hypothetical formalized language of science a body of theory titled "causal theory of reference" <u>is</u> just adding more <u>theory</u>. But Skolem's argument, and our extensions of it, are not affected by enlarging the theory. Indeed, you can even take the theory to consist of <u>all true sentences</u>, and

> there will be many models - models differing on the
> extension of every term not fixed by OP (or whatever
> you take OP to be in a given context) - which satisfy
> the entire theory. If 'refers' can be defined in terms
> of some causal predicate or predicates in the metalang-
> uage of our theory, then, since each model of the object
> language extends in an obvious way to a corresponding
> model of the metalanguage, it will turn out that <u>in
> each model M</u>, <u>reference</u> is definable in terms of <u>causes</u>$_M$;
> but, unless the word 'causes' (or whatever the causal
> predicate may be) is already glued to one relation with
> metaphysical glue, this does not fix a determinate exten-
> sion for 'refers' at all.   9

Remarks at the end of the paper throw further light on these points

and also help us to situate the "model-theoretic" argument in relation

to the "sophisticated realism" of "Reference and Understanding". Putnam

claims that the "model-theoretic" argument arises only to the extent

that we both (a) explicate understanding a language in terms of the

ability to <u>use</u> the language, and (b) ask "what the possible 'models'

for the language [are], thinking of the models as existing 'out there',

<u>independent of any description</u>." Since the understanding of the language

must determine the reference of the terms, given context of use, the

identification of understanding with use implies that use must deter-

mine reference. But, in accepting (a) and (b), we commit ourselves to

the view that "the language has a full program of use, but it still

lacks an interpretation," and thereby to a problem "which <u>can</u> only have

crazy solutions": "Either the use <u>already</u> fixes the interpretation

or <u>nothing</u> can." To appeal to a "causal theory" of reference at this

point is to hope "that the <u>world</u> will pick one definite extension for

each of our terms even if <u>we</u> cannot." In other words, only if THE WORLD

itself selects one of the possible "correspondences" consistent with

our use as the <u>true</u> relation of reference can a "causal theory" of

reference enable the externalist to evade the problems posed by the "model-theoretic" argument.[10]

A number of commentators have questioned why, if the "model-theo-retic" argument works for an "Ideal" theory, it should not work equally well for _any_ consistent theory. Will not any consistent theory (of an appropriate cardinality) permit a SAT interpretation of its language into THE WORLD? Clark Glymour, for example, claims that "a perfectly similar argument holds for _non-ideal_ theories that are, or have been, rationally acceptable," and argues that Putnam must either (a) hold that, e.g. Newtonian mechanics, _was_ true but is no longer so, or (b) hold that no non-ideal theories are ever rationally acceptable.[11] David Lewis has also voiced a measure of perplexity on this question.[12] The answer, however, is relatively straightforward, and is presented fairly explicitly in "Models and Reality". The contention that an "Ideal" theory cannot intelligibly be supposed to be false is not based merely upon the existence of a SAT interpretation which makes all sentences of the theory come out true. It is also required that such a SAT interpret-ation of the theory _meet all operational and theoretical constraints._ This is crucial because it removes obvious grounds for maintaining that, in spite of the existence of the SAT interpretation, the theory _might still be false_ - that is, _operationally_ false. But this require-ment is not satisfied in the case of non-ideal theories:

> The theory T is, we may suppose, well confirmed at the
> present time, and hence rationally acceptable on the
> evidence we _now_ have; but there is a clear sense in which
> it may be false. Indeed, it may well lead to false pre-
> dictions, and thus conflict with [the set of operational
> constraints]. But T_I, by hypothesis, does not lead to any
> false predictions. Still, the metaphysical realist
> claims,...T_I may be, in reality, false.   [13]

It is the inability to make good on this claim that a theory might be false "in the face of all possible justification" that renders metaphysical realism incoherent, according to Putnam. But no such difficulties attend the claim that a non-ideal theory might be false.

A related misunderstanding of Putnam's argument occurs in a paper by Gregory Currie[14]. Currie argues that, given the resources that Putnam utilises in establishing that an "Ideal" theory must be true, one can construct a parallel argument establishing that such a theory must be false. The strategy is to proceed exactly as Putnam does, but to select an interpretation (let us call it NON-SAT) under which at least one non-observational sentence of the theory is false. The availability of such an interpretation (and, indeed, of a plethora of the same) follows from precisely the same model-theoretic results cited by Putnam. Intuitively, all we would require is a SAT interpretation for the theory T^*, where the latter differs from the "ideal" theory T_I only in having not-S where T_I has S, for some non-observational S, while remaining consistent. Currie maintains that NON-SAT cannot be excluded as an "unintended" interpretation by imposing the requirement that an acceptable interpretation not make the theory come out false. For, he claims, this would be to rule out by fiat the very possibility that the externalist wishes to entertain. He concludes that what Putnam's line of reasoning actually establishes is that "there is a range of interpretations consistent with all acceptable constraints and that [the "Ideal" theory] is true in some of these and false in others!" There is nothing contradictory in this, he maintains, since it is not being claimed that T_I is both true and false under the same interpretation. But it does follow that theoretical sentences are neither true nor false in any interpretation-inde-

pendent sense. Thus what Putnam's argument actually supports is "a trad-
itional and familiar version of scientific instrumentalism."

What Currie misses is the role played in Putnam's argument by our
general intention to refer in such a way that our beliefs come out _true_.
This intention is not something that the externalist can deny, nor, one
would think, is it something the externalist would have any wish to deny.
To acknowledge this intention is not to grant the anti-realist the "un-
reasonable" premiss that a theory cannot come out false on an acceptable
interpretation (or that an "Ideal" theory cannot come out false on an
acceptable interpretation). But it has the same force as this premiss
if we also grant Putnam's contention that there is nothing _else_ in the
externalist model that can ground the sort of determinate reference
relation required if we are to intelligibly entertain the possibility
that an "ideal" theory might be false. For, given this contention,
there will be nothing that can serve to validate _some_ interpretations
of an "Ideal" theory - the SAT interpretations - as "intended" _except_
our intention that our beliefs be true. Currie is quite right when he
claims that both SAT and NON-SAT interpretations of the "Ideal" theory
will satisfy all operational constraints. But the "theoretical" con-
straint that we intend to "get it right" will serve to exclude the
NON-SAT interpretations as "unintended", since it is the only constraint
that effects a distinction between otherwise acceptable interpretations
in the case of an "Ideal" theory.

V-b   How Not to OutSmart the Anti-Realist

A popular strategy, in debates between proponents of externalist
and non-externalist theories of truth, is the appeal to that paradigm
of analytical philosophy, the counter-example. It will be claimed,
usually by a philosopher of the externalist persuasion, that a proposed
account of truth must be rejected because it cannot do justice to the
truth (falsehood), or the possible truth (falsehood), of a particular
statement S. Putnam himself employs just such a strategy, albeit in a
more schematic manner, in "Reference and Understanding", where he ad-
vances the "Idealistic Fallacy" argument for a realist account of
truth. The contention is that, for any proposed non-realist explication
of truth in terms of a property $\emptyset$, it makes sense to say of a statement
S that S possesses $\emptyset$ yet may still not be __true__. In the present section,
I shall examine the use of the "counter-example" strategy in the inter-
ests of rebutting the "model-theoretic" argument. I shall take, as rep-
resentative examples of this theme, papers by J.J.C. Smart[15] and Alan
H. Goldman[16]. In discussing Smart's paper, I shall also address a sugg-
estion by Peter Smith as to how Putnam might respond to such criticism
of the "model-theoretic" argument.[17]

Smart asks us to consider the following case. We have two theories,
T_I and T_I^*, which agree in their descriptions of our four-dimensional
space-time manifold, and which satisfy all possible operational con-
straints relative to that manifold. The theories differ only in that T_I^*

> asserts that our space-time manifold is a cross-section
> of a five-dimensional manifold which contains as another
> cross-section another space-time manifold, also, like
> ours, containing galaxies, stars, planets, and so on.
> Let us suppose that according to T_I^* we can never know
> about this other space-time sub-universe because this

sub-universe and our own are precluded by the laws of
nature from causally interacting with one another.   18

Smart argues that, given the theoretical constraint of simplicity,
it would never be rational to prefer $T_I{}^*$ to T_I, but that, nonetheless,
$T_I{}^*$ might be <u>true</u>, since the universe might be less simple than we
anticipate.

Smith questions whether this line of reasoning can establish the
possible <u>falsity</u> of an "Ideal" theory. What facts does T_I <u>get wrong</u> if
$T_I{}^*$ is true? The obvious answer is that T_I mischaracterises the dimen-
sionality of the physical universe. But, Smith argues, we cannot give
any content to the claim, in $T_I{}^*$, that the separate four-dimensional
sub-universes are "genuinely embedded in a physically real five-
dimensional manifold" if, as $T_I{}^*$ also maintains, the sub-universes are
causally isolated from one another. And if we allow for causal inter-
action between the sub-universes, the "empirical ideality" of T_I
becomes moot. If, on the other hand, we pare down Smart's example to
the "key idea" that there may be causally unrelated sub-universes (or,
more generally, "empirically inaccessible regions of the world"), then,
in so far as T_I observes the plausible theoretical constraint that a
theory should eschew "surplus content", our "Ideal" theory, T_I, will
refrain from making <u>any</u> claims about the existence or nature of other
sub-universes. In this case, the sort of possibility envisaged by
Smart will not demonstrate that T_I might be <u>false</u>, but only that it
might be <u>incomplete</u>, Smith concludes.

Can this strategy rescue Putnam from Smart's objection? <u>Prima
facie</u>, it seems that it can. Suppose we term the claim that belonging
to an "Ideal" theory is <u>sufficient</u> for the truth of a statement the

"weak" thesis, contrasting this with the "strong" thesis that belonging to an "Ideal" theory is both sufficient and _necessary_ for the truth of a statement. Then Smith's claim is that Smart's example may count against the strong thesis but not against the weak thesis. Since Putnam's argument against externalism rests upon (a) the premiss that the latter requires the denial of the weak thesis, and (b) the "model-theoretic" argument _for_ the weak thesis, Smith's claim, if correct, would seem to completely defuse Smart's example as an objection to Putnam. But can Putnam afford to grant the force of Smart's example against the _strong_ thesis? If we recall the details of the "model-theoretic" argument, I think it will be apparent that he cannot do so.

For to grant the force of Smart's example against the strong thesis is to allow that certain statements might be true although not contained in the "Ideal" theory. The obvious question now, however, is, in virtue of _what_ are such statements true? The truth of statements which _are_ contained in the "Ideal" theory is given, on Putnam's account, by the SAT interpretations of the language. Let S represent a statement such that: (i) S is true, and (ii) neither S nor not-S is contained in the "Ideal" theory T_I. There is, indeed, an interpretation of $T_I \cup S$, call it SAT*, on which both T_I and S come out true. But there is also an interpretation of $T_I \cup$ not-S, call it SAT**, on which T_I comes out true and S comes out false. But since, _ex hypothesi_, we remain agnostic about S in our "Ideal" theory, our intention that our beliefs come out true cannot select SAT* over SAT** as the "intended" interpretation of our language. Thus the truth of S cannot be explained in the same way as the truth of those statements contained in the "Ideal" theory. It seems, in fact, that we must say that, if S has a determinate truth-

value, it is in virtue of its corresponding to "the way THE WORLD is", independently of whatever SAT relations may obtain. And this requires that there be a unique and determinate relation of reference between at least some of the terms of the language - those employed in the formulation of S $\rightarrow$ and entities, or sets of entities, in THE WORLD. But, if such a relation exists and is a ground for the truth of some statements expressible in the language, we must surely draw the further conclusion that it is this relation, and not SAT, that is the genuine "intended" interpretation of the language as a whole. And then there is no reason why statements contained in the "Ideal" theory itself should not be false on this interpretation, even if true on SAT. Thus Putnam cannot grant the force of Smart's objection against the strong thesis. For, if he allows that an "Ideal" theory may be incomplete with respect to certain expressible statements, then he must also allow that it may be false.

I have argued that Putnam cannot avail himself of Smith's proposed strategy to defend himself against Smart. But this should give only momentary comfort to the cause of externalism. For, as may already be apparent, the very considerations which prevent Putnam from ceding the strong thesis to Smart also provide him with an answer to Smart's criticism. Smart believes that he can defend metaphysical realism against the "model-theoretic" argument by producing a counter-example to the argument's conclusion, rather than by directly addressing the argument itself. But the counter-example can be made out only if Putnam's argument is flawed. To simply offer the "counter-example" as a refutation of the argument is thus to beg the question at issue.

For consider what is involved in Smart's claim that T_I^* might be

true. He maintains that the conceptual resources of mathematics and physics will permit us to describe the five-dimensional manifold postulated by the theory (or, perhaps, the alternate sub-universes which the theory envisages). But the truth-values of the sentences in T_I^* which incorporate these descriptions depend, according to the externalist, upon referential relationships that obtain between the relevant terms, on the one hand, and entities in THE WORLD, on the other. Such sentences can have determinate truth-values only if the relevant terms uniquely and determinately refer. But the "model-theoretic" argument purports to show that there is nothing in the externalist picture that can fix the referents of terms in a language, save for the considerations adduced in the characterisation of SAT, and this will apply to the terms employed in T_I^* as much as to any other terms. If Putnam's argument is correct, then the cosmological claims in T_I^* can be supposed to have a determinate truth-value only by way of a SAT interpretation. But SAT presupposes the intention that one's beliefs be true, and this, in turn, requires that one has the beliefs in question. And, ex hypothesi, members of our "Ideal" community of inquirers would not embrace the cosmological speculations of T_I^* - in fact, as Smart points out, the rational belief, now or under "Ideal" circumstances, is that these speculations are false. So, if statements can have a determinate truth-value only by way of a SAT interpretation, the speculations of T_I^*, if they have a determinate truth-value at all, will be false. Such speculations can be supposed to be true only if Putnam's strictures upon the possibility of SAT-independent determinate reference within the externalist framework are rejected. As we shall shortly see, there are those who have argued that these strictures

should be rejected. But, until such arguments are made out, Smart's "counter-example", and, indeed, any other example of that sort, is surely question-begging. And, one might add, if such arguments are made out, Smart's example and others like it will be unnecessary;·for, if there are SAT-independent constraints on reference, as other philosophers have argued, then any claim in an "Ideal" theory that goes beyond what is strictly dictated by operational constraints may be false

Alan H. Goldman maintains that defenders of externalism have mistakenly tried to engage their opponents at the level of "global error", where what is at issue is the possibility·of our being massively deceived in our beliefs. Standard examples of purported global error - "brain in a vat" and "evil demon" scenarios - invite the response that the supposedly false beliefs are actually true when the believer's language is given its proper interpretation. This, as we have seen, is the sort of strategy adopted by Putnam in both the "brain in a vat" and the "model-theoretic" arguments. Goldman maintains that the realist should begin with a case of "local error" - a situation where we have the possibility of undetectable error in respect of a single belief - and then argue from this to the possibility of global error. He offers a hypothetical example of local error which purportedly exhibits the following qualities: (a) a belief B is prima facie false; (b) B satisfies all relevant criteria of assertibility, so that the falsity of B cannot be explained in terms of. a non-realist notion of truth; and (c) it is not possible to make B come out true by means of the sorts of anti-realist strategies standardly employed to counter purported

examples of the possibility of global-error. I shall grant Goldman
that, if it can be shown that we are committed to the possibility of
local error that resists an anti-realist explanation, then an analogous
conclusion can be established in the case of global error. Indeed, I
made a related claim in my criticism of Smith's response to Smart[19].
And even if we couldn't move beyond the local case, this would still
suffice to rebut Putnam's claim that an "Ideal" theory could not be
false, to the extent that B would be preserved within that theory.
I shall argue, however, that Goldman is wrong in claiming that his
"fanciful" local case satisfies condition (c) above. In so doing, I
shall attempt to curb his realist flight of fancy before it leaves the
ground.

Goldman offers the following example, which is supposed to illus-
trate the possibility that there might be "a belief that remains for-
ever consistent with others and immune from falsification, but that is
false:"

> Imagine...a person S and his physical replica R, an
> atom for atom duplicate. In this story shortly after R
> materialises, S dematerialises without a trace. R's
> seeming memories now match S's former genuine memories.
> Specifically, R remembers having had an unhappy child-
> hood, which, of course, he never had. It appears, never-
> theless, that his belief must meet all pragmatic and
> coherentist tests of truth: it can be verified but not
> falsified. Complete evidence here will fail to entail
> truth; it appears to be the realist's notion of corres-
> pondence, or lack of it, that determines the falsity
> of R's belief.   [20]

Goldman maintains that the anti-realist cannot argue that the falsity
of R's belief consists in its failure to cohere with <u>our</u> understanding
of the situation, because this would not account for the fact that R's

belief is false "in his world" and not simply false in relation to ours. He further maintains that the anti-realist cannot appeal to some test or procedure that would serve to distinguish between S and R. I think Goldman's argument for this point is unsound, but I shall not argue this here. Rather, I shall take up the claim that the "local" case is able to circumvent standard anti-realist counters to "global case" arguments for realism.

One such counter is that the supposed victims of global error - the permanently envatted brains or the subjects of demonic deception - can be taken to "occupy their own worlds, those defined by their sets of beliefs...[within which] their beliefs could be held true when coherent or verified, and true of the objects in those worlds as their beliefs define them."[21] Goldman claims that no such strategy will work in the local case":

> We cannot assign phenomenal referents here, since the reference of terms like 'childhood experience' will be fixed by uses of such terms to refer successfully by other members of R's linguistic community. His belief cannot be taken to refer to his imaginary experience, since there will be a distinction in his world, as in ours, between imaginary and genuine experiences and 'childhoods. We cannot then map the belief onto an object that would make it come out true, despite the belief's meeting all operational and verificationist tests. [22]

Because R "intends to use his terms as others use them", it is possible - given Putnam's "linguistic division of labour" - to make sense, in the local case, of the idea that "the mapping that makes the belief or set of beliefs true may not capture their intended referents... because reference is determined by practice outside the local belief or theory that borrows and assumes the fixation of referents for its

terms."[23]

Note that, in the quoted passages, Goldman seems to equate (i) assigning phenomenal referents to the terms of a language, and (ii) mapping that language onto objects,in such a way that beliefs come out true. Doing the first of these things is one way of doing the second, but it is not the only way. I think that Goldman is conflating two distinct arguments that Putnam has directed against realist contentions.[24] In rejecting the claim that we might be "brains in a vat", Putnam argues, as we have seen, that, given the externalist framework, plausible constraints on reference, and our "intuitive" sense of an "intended" interpretation, we should take the referents of terms in vat-English to be things "in the image" (or elements in the causal process productive of such things): the strategy, here, is to make the brain's belief that it is not a brain in a vat come out true by assigning "phenomenal" referents. But this strategy will not always work, so long as we retain our intuitive notion of an "intended" interpretation, and I think Goldman is right in claiming that this strategy cannot be employed to counter his local case. For, if we assign "phenomenal" referents to R's terms while not doing this for lexically identical terms employed by members of the linguistic community to which R belongs (henceforth LC), we will not be respecting our intuitions about "intended" interpretations, given R's intention to conform to the practice of LC in his use of the language.

As we have already noted in our discussion of Smart, however, a crucial feature of the "model-theoretic" argument is the claim that we cannot preserve our intuitive notion of an "intended" interpretation

within the externalist framework. The only thing that can pick out certain of the interpretations satisfying all operational and theoretical constraints as "intended" is our intention to refer in such a way that our beliefs come out _true_, coupled with the availability of SAT interpretations that _make_ all of our beliefs come out true. Goldman's claim that we cannot map R's belief onto objects in such a way as to make that belief come out true is, then, the claim that the anti-realist cannot avail himself of a SAT-like interpretation incorporating R's belief that he had an unhappy childhood. And the _argument_ for this claim is that no interpretation that makes this belief come out true can count as "intended" because it will fail to respect those extra-local features of practice that operate as constraints on the reference of terms in R's language. Given those constraints, Goldman argues, all acceptable interpretations must entail that R's belief is _false_ "in his world".

But what is "R's world" in this context? _Not_ a solipsistic world, or a world defined by R's own beliefs; for, as Goldman stresses in rejecting the idea that R's terms can be ascribed "phenomenal" referents, R intends that his linguistic practice conform to that of other members of LC. "R's world" is, rather, the world he shares with the other members of LC in Goldman's fanciful universe, and R's belief that he had an unhappy childhood is merely a first-person (for R) formulation of the communal belief that R had an unhappy childhood. To ask, "Can the anti-realist answer Goldman by offering a SAT-like interpretation that makes R's belief come out true?", is; then, to ask whether we can produce a SAT-like interpretation of the maximally coherent beliefs of LC, including the belief that R had an unhappy

childhood, that cannot be rejected as "unintended". But, given that
(i) such a SAT-like interpretation will treat the belief that R had an
unhappy childhood uniformly with the other beliefs of LC in its ascrip-
tion of referents to the component terms of the corresponding sentences,
and (ii) R's intention that his usage conform with that of other members
of LC allows us to identify the assertoric content of his belief that
he had an unhappy childhood with that of the communal belief that R had
an unhappy childhood, there are no obvious grounds upon which Goldman
could argue that such a SAT-like interpretation fails to respect extra-
local constraints on reference and is therefore "unintended". And on
the SAT-like interpretation, R's belief is _true_ in his world, in which
case the "counter-example" fails to demonstrate the need for realist
notions of truth and falsity.

At times, Goldman seems to be suggesting that what rules out such
a SAT-like interpretation is that we are _given_ the correct, or "intend-
ed", interpretation of the language of LC, and thereby the falsity of
R's belief about his childhood, in the very characterisation of the
example. Thus we find him saying that "in developing the possible world
as we choose, we have special access to the correspondence relations
between beliefs and facts within it."[25] But Putnam's "model-theoretic"
argument, as we have stressed, purports to show that there are no SAT-
independent constraints available to the externalist that can serve
to fix a "unique interpretation" of a language, To try to defend realism
against Putnam by offering an example that _specifies_ such an interpret-
ation by _fiat_ makes sense only if we countenance the existence of some
intelligence that stands to actual linguistic practice in the same

relation that we, as tellers of the story, stand in to the practice of members of LC. In the absence of a suitable argument for the existence of such an Ideal Referer, an example of "realist truth" that relies upon such a specification of an "intended" interpretation serves only to widen, rather than narrow, the gap between realist truth and our actual linguistic practice.

V-c Models and Ideality

In the remainder of this chapter, I shall critically examine responses to the "model-theoretic" argument that address themselves, correctly I think, to the argument itself, rather than offering purported counter-examples to the argument's conclusion. It will be helpful, in this context, to separate two stages in the argument, and to consider critical responses to each of these stages in turn. The first stage of the argument - steps (a) and (b) in the outline of the argument at the beginning of sub-section 3, section V-a, above - claims to demonstrate that operational and theoretical constraints cannot serve to exclude, as "unintended", SAT interpretations of the language of an "Ideal" theory on which that theory comes out true. The second stage of the argument - steps (c) and (d) in the outline - is intended to counter the suggestion that there are any other constraints available to the externalist that can perform this function, and to establish, thereby, that an "Ideal" theory could not fail to be true. Let me, in uninspired fashion, refer to these two stages as Stage I and Stage II respectively. In the present section, I shall briefly discuss certain critical responses to Stage I of the argument. In the following two sections, I shall examine in more detail critical responses to Stage II.

Philosophers who have challenged Stage I of the "model-theoretic" argument have tended to do so on one or other of the following grounds: (i) it is claimed that there are technical difficulties attending Putnam's appeals to results in model theory; (ii) it is claimed that the notion of an "Ideal" theory, as employed by Putnam in the argument, is unserviceable. Challenges of the first sort have come from critics who

have assumed - perhaps because of the framework employed by Putnam in the "Models and Reality" presentation of his argument - that the argument requires the Lowenheim-Skolem theorems in model theory. Such critics have argued that the Lowenheim-Skolem theorems are not valid for second- and higher-order languages, and that Putnam's argument therefore rests upon the tacit, and unjustified, assumption that the "Ideal" theory, incorporating the necessary operational and theoretical constraints, could be fully formalised in a first-order language Pearce and Rantala, for example, charge that "it is an ad hoc move... for Putnam to assume without further argument that all theoretical constraints must be formalised in first order logic."[26] And Ian Hacking maintains that all that Putnam has shown is that "you cannot succeed in reference by stating a set of truths expressed in first-order logic", and that there is no reason to think that we can give a first-order formulation of the language of physics.[27]

As a number of other writers have pointed out, however, the Lowenheim-Skolem theorems play no essential role in Putnam's argument against metaphysical realism, although they are required to generate the "Skolem Paradox" with which the discussion of "Models and Reality" begins. The general import of the Lowenheim-Skolem theorems is that a set of first-order sentences that has an infinite model will have models of every infinite cardinality. What this implies, as far as Putnam's argument is concerned, is that, if the "Ideal" theory is first-order and has some infinite model, then it cannot determine its models up to isomorphism.[29] That Putnam's case against the metaphysical realist doesn't require the Lowenheim-Skolem theorems becomes apparent when we realise that the realist's discomfort is not assuaged even if we assume

that the "Ideal" theory can single out an isomorphism class of models, for the extensions of individual terms in the language admit of permutations within that class:

> The general point is that the extensions of the terms of the language are not fixed within a class of isomorphic models. Permuting the domain of a model results in a change of the extensions of the terms of a theory, but since a permutation is simply a one-to-one mapping of every element of the domain onto another, isomorphism is preserved. In Putnam's argument this means that all terms whose extensions are not fixed by the valuation OP will be indeterminate within a class of isomorphic models. 30

Indeed, we may recall that Putnam's later variations on the "model-theoretic" theme - the "cats and cherries" argument in RTH and the "parable" about male and female semantics in the Introduction to Realism and Reason - turn upon the possibility of precisely this kind of permutation of the extensions of the terms within a given domain. The model-theoretic resources upon which such arguments draw - essentially nothing over and above the completeness theorem - are modest, and do not raise any of the sorts of technical difficulties sketched above.

The second type of challenge to Stage I of the "model-theoretic" argument questions whether there is an intelligible notion of "Ideal Theory" that can serve Putnam's purposes. Michael Devitt[31], for example, argues that the notion of a theory that satisfies "all operational constraints" is unclear in that we are not told whether "all possible observational evidence" is to include not only the results of passive observation at every point in space-time, but also the observed results of every possible experiment at every such point. And, if experiments

be permitted, which instruments are we allowed to employ? The obvious response to this line of criticism is to admit the unclarity, but to maintain that it does not blunt the force of the anti-realist argument because, however we proceed in precising the notion of an "Ideal" theory, the realist will still be committed to the existence of a conceptual gap between the deliverances of such a theory - which satisfy all criteria of "warranted assertibility" - and the truth.[32] But Devitt seems to have an answer to this response:

> Suppose it is the case that each aspect of reality is
> capable of playing some causal role. Then perhaps that
> aspect would be detectable by humans using some possible
> instrument. So according to my epistemology, on some
> very liberal interpretations of "possible evidence"
> [the "Ideal" theory] could not be false. But that is
> no reflection on my realism.   [33]

This answer doesn't work, however, for it fails to bridge the conceptual gap between "Ideal" assertibility and externalist truth. The realist must regard it as a contingent matter that each aspect of reality is detectable, if this is indeed the case, and is therefore committed to the existence of possible worlds in which there exist undetectable aspects of reality capable of rendering false a theory which is "Ideal" for the denizens of those worlds. But the realist is entitled to uphold the existence of such possible worlds only if he can show how such falsity is possible in the face of the "model-theoretic" argument.

V-d The "Canonical" Strategy .

Most critics of the "model-theoretic" argument have been prepared to grant Stage I of the argument, but have balked at Stage II on the grounds that there are further constraints - of a loosely "causal" nature - that serve to determine reference in a (metaphysical) realistically acceptable way. The "canonical" realist strategy, as we may term it, is to assume that, in David Lewis's words, "there must be some additional constraint on reference: some constraint that might, if we are unlucky in our theorising, eliminate all the allegedly intended interpretations that make the theory come true."[34] Exponents of this strategy have tended to argue in the following manner:

(i) Putnam assumes that the only plausible constraints on reference available to the realist are those imposed by the acceptance of an uninterpreted, or partially interpreted, set of sentences.

(ii) This assumption is incorrect: our use of language as part of our more general commerce with one another and with the world constrains the reference of our terms, and the realist can appeal to such constraints in explaining how determinate reference is possible and why SAT-like interpretations are not acceptable construals of reference.

(iii) What is especially puzzling about Putnam's argument is that he himself, in his "pre-critical" writings, was among the foremost proponents of the view that reference is fixed, not by the sets of sentences held true by the speakers of a language, but by causal relations between speakers and their environment.

Let me briefly illustrate the prevalence of variations upon these themes in the critical literature before turning to an assessment of this sort of strategy.

In his already cited paper, "Putnam's Paradox", David Lewis attributes to Putnam the view that the only alternative to "magical" theories of reference is what Lewis terms "global descriptivism" - the thesis that the "intended" interpretation of a language is the one that makes the speakers' "total theory" come out true, or the one that most closely approximates that ideal. "Global descriptivism" is a generalisation of the "descriptions theory" of reference more familiarly offered as an analysis of the mechanisms determining the reference of particular classes of terms, such as proper names and natural kind terms. According to Lewis, Putnam espouses "global descriptivism" because he believes that "constraints that work within it are the only possible constraints on reference... [because] global descriptivism is imperialistic: it will annex any satisfactory alternative accounts of constraints on reference."[35] This, Lewis maintains, is why Putnam dismisses the suggestion that the realist can evade the "model-theoretic" argument by appealing to a "causal theory" of reference. Alvin Goldman, citing a paper by Alvin Plantinga[36], offers a similar diagnosis of Putnam's views on the determination of reference:

> Putnam seems to suppose that the terms of our language
> get their meaning or extensions, somehow, by virtue
> (solely) of the set theoretical models of first order
> formalisations of the body of our beliefs. In other
> words, roughly, our terms get their meaning and exten-
> sions by virtue of a vast network of implicit defin-
> itions. [37]

And, in a more technically sophisticated vein, William Demopoulos maintains that the entire burden of the "model-theoretic" argument rests upon what he terms "the doctrine of partial interpretation" -

the doctrine that the reference of the theoretical vocabulary of a theory T which is "partially interpreted" through a specification of reference for the O-vocabulary "is fixed by the formal theoretical sentences of T."[38]

Exponents of the "canonical" strategy are virtually unanimous in claiming that Putnam unjustifiably restricts possible constraints on interpretation in the ways just suggested, although they differ both in their willingness to identify additional constraints and in the kinds of additional constraints that they propose. Demopoulos, for example, cites "a strong pre-analytic intuition that our linguistic practice must incorporate constraints that go beyond the mere satisfaction of a set of formal sentences", but refrains from further speculation on the grounds that "the notion of a linguistic practice - of an established use - is hardly well understood."[39] Ian Hacking, on the other hand, is less circumspect, affirming his belief that

> assuring reference is not primarily a matter of uttering
> truths, but of interacting with the world...Language is
> embedded in a wide range of doings in the world...Cher-
> ries are for eating, cats, perhaps, for stroking. Once
> speech becomes embedded in action, talk of Lowenheim and
> Skolem seems scholastic...We can do very little with
> large numbers except talk about them. With cats we relate
> in other ways than speech.   [40]

As to the nature of the feature of our "interacting with the world" that establishes a determinate relation of reference between terms in our language and particular things in the world, most writers have taken the relevant constraints to consist in the existence of appropriate causal relationships between terms and their referents. Clark Glymour, for example, speaks of "causally determined reference relat-

ions": "Roughly, our physical and social circumstances, and sometimes perhaps our beliefs as well, determine together a series of links connecting words and objects, and thus delimiting the admissible interpretations of our theories."[41] Michael Devitt maintains that the realist should adhere to a "causal theory" of reference, and that such a theory allows the realist to view even an "Ideal" theory as something that arises

> out of a causal interaction between the believers and a
> reality independent of those beliefs. Related to this it
> will always seem intelligible to the realist to suppose
> that the theory might not be true. The realist will also
> see the reference of the terms in that theory as basic-
> ally, and in general, determined by these causal inter-
> actions; at least he will if he combines a causal
> theory of reference with his realism.   42

An alternative approach takes the necessary additional constraints on reference to be given _not_ by a "causal theory" of reference, but, rather, by the existence of "structure" in the world itself. G.H. Merrill, for example, argues that Stage I of Putnam's argument rests upon the assumption that the world can be adequately represented as a set of objects, whereas realists "typically hold that not only are there objectively existing entities (both observable and unobservable) in the world, but also that these entities' bear to one another certain objective relations": if we are to capture the realist conception of the world in our semantics, we must represent the world as "a set together with a class of relations among the members of that set":

> The world is (or corresponds to) a structured domain
> whose constituents are independent of any particular
> representation we have of it. The realist need not claim
> to know the make-up of either the domain of the world
> or the relations on it. He asserts only that the world

> is (or corresponds to) such a structured domain, though
> which one it is he does not (even cannot) know.   43

Merrill claims that the additional constraints imposed by a "structured
domain" semantics enable the realist to give sense to the hypothesis
that an "Ideal" theory might be false without having to appeal to the
notion of an "intended" interpretation (the notion that the "causal
theory" approach seeks to rehabilitate). An acceptable interpretation
of the language of the "Ideal" theory must respect the intrinsic
structuring of the world, and Putnam's argument will go through only
if there is a structure-preserving interpretation of that language in
the world under which every sentence of the "Ideal" theory comes out
true. But there may fail to be such an interpretation: "Given two
structured domains (the world and that imposed by the model [of the
"Ideal" theory] on its universe) it is not always possible to 'force'
them into isomorphism."[44] David Lewis has endorsed an approach similar
to that of Merrill, maintaining that not all things and classes are
properly treated as equal when the selection of reference-classes or
terms in our language is at issue:

> Only an elite minority are carved at the joints, so that
> their boundaries are established by objective sameness
> and difference in nature. Only these elite things and
> classes are eligible to serve as referents...When we
> limit ourselves to the eligible interpretations, the
> ones that respect the objective joints in nature, there
> is no longer any guarantee that (almost) any world can
> satisfy (almost) any theory.   45

That our use of language occurs in the context of our more general
intercourse with non-linguistic reality seems obvious. No less obvious,
it would appear, is that such features of language use can operate as

constraints upon the interpretation of the terms in a language. The contention of Putnam's critics that Putnam overlooks such possible resources for the realist therefore invites the question how he could have overlooked something so transparent. Lewis's explanation, as we saw above, is that Putnam is committed to "global descriptivism" and believes that the latter can imperialistically annex any further constraints that might be proposed. But this explanation raises puzzles of its own, given that Putnam's "pre-critical" writings contain some of the strongest attacks upon the traditional "descriptivist" programme in the philosophy of language. Putnam uses his "Twin Earth" examples to argue that the primary determinant of the reference of a term in the idiolect of a community of speakers is the nature of the "stuff" in the world that is picked out, in an indexical, or quasi-indexical manner, by the linguistic practice of that community. Far from its being the case that the referent is selected in virtue of its satisfying a set of associated descriptions (or the majority of such a set), the referent of a term may in fact <u>fail</u> to satisfy such a set (or the majority thereof) on Putnam's "pre-critical" view of things.[46] In the light of this, it is difficult to accept, without strong textual evidence, the thesis that Putnam ignores the possibility of "causal" constraints on interpretation because he is blinded by his commitment to "global descriptivism".

The foregoing anomaly confronting the "canonical" strategy has not gone completely unremarked by Putnam's critics, and at least one recent commentator has ventured a more extensive treatment of the issue. It will be helpful to provide a summary of the explanation proposed by the latter writer - Carsten Hansen, in the paper "Putnam's Indeterminacy

Argument"[47] - for it may serve as a suitable foil for a subsequent

assessment of the "canonical" strategy in general. Hansen focuses upon

Putnam's response, in the Introduction to <u>Realism and Reason</u>, to the

charge that the "model-theoretic" argument ignores the possibility

that reference is fixed by "something non-psychological." Putnam's

answer is that

> the idea that the "non-psychological" fixes reference -
> i.e. that <u>nature itself</u> determines what our words stand
> for - is totally unintelligible...To say in one's most
> intimidating tone of voice "I believe that causal conn-
> ections determine what our words correspond to" is
> only to say that one believes in a <u>one-knows-not-what</u>
> which solves our problem <u>one-knows-not-how.</u>   [48]

What is puzzling about this answer, Hansen claims, is why Putnam wants

to deny the realist the sort of strategy he himself adopted in such

papers as "The Meaning of 'Meaning'" - the strategy of explicating

reference, not in terms of "what is in the head", but in terms of

indexicality and the "linguistic division of labour":

> He completely ignores the possibility of appealing to
> the role of the linguistic community and the phenomenon
> of indexicality in determining the interpretation of the
> language - at least when discussing what he now considers
> to be legitimate realist options. It is also remarkable
> that causal connections are rendered completely inacc-
> essible in Putnam's account - an appeal to them would be
> an appeal to a "<u>one-knows-not-what</u> which solves our
> problem <u>one-knows-not-how.</u>"   [49]

Hansen's explanation of Putnam's conduct is that the "model-theo-

retic" argument tacitly rests upon what he terms "a Lockean epistemol-

ogical premiss." He attributes to Putnam the assumption that any ade-

quate realist answer to the "model-theoretic" challenge must satisfy

two conditions:

> (1) What determines the interpretation of a term can only
> depend on circumstances known by the linguistic community
> and therefore in principle accessible to each individual
> member....
> (2) We have knowledge of, or access to, (mental) rep-
> resentations only.   50

Furthermore, we can explain why Putnam denies the realist use of the

"Meaning of 'Meaning'" strategy only if we also ascribe to Putnam a

radical reading of the second of these conditions:

> The representations to which we have access...are in a
> significant respect like Lockean "ideas" - they are
> mental entities of a sort that form a "veil" between us
> and the external world. Only on such a construal would
> it be plausible to claim that an exhaustive division _
> between the 'psychological' and the 'non-psychological'
> could be drawn with the 'psychological' understood as
> something to which we have direct access, and the 'non-
> psychological' as something to which we, in effect,
> have no access   51

Given this "Lockean" reading of condition (2), Putnam's off-hand dis-

missal of realist appeals to "causal connections" and "non-linguistic

facts" becomes intelligible, for "it will be question-begging to assume

the possibility of having knowledge of the external, non-representational

objects in terms of which the language is to be interpreted."[52] The

ascription of a "Lockean" reading of condition (2) to Putnam is further

justified, Hansen maintains, by the fact that it is necessary if we are

to account for the terminology that Putnam employs in characterising

the realist's difficulties - his talk of the need to single out "a

determinate correspondence between words or mental representations and

external objects," and to grasp "what is outside the mind." But if the

"model-theoretic" argument does require such a "Lockean" epistemologic-

al premiss, then, Hansen argues, it has little force against metaphys-

ical realism (or, more precisely, against what Hansen terms "minimal realism" see below): for the realist is not committed to such a premiss, nor is such a premiss independently plausible. Thus it is open to the realist to reject the "Lockean"-assumption and offer the sort of "causal" story about reference that is the core of the "canonical" strategy.

Our examination of Putnam's writings in earlier chapters of the present work should, I think, render us somewhat skeptical as to the claim that he subscribes to Hansen's "Lockean" premiss, and equally skeptical in the absence of textual corroboration, of the claim that he assumes that the <u>realist</u> must subscribe to such a premiss. It is <u>Kant</u>, and not Locke, we may recall, that Putnam views as his philosophical precursor, and it would surely be a mistake to construe Kantian "idealism" as a form of skeptical Lockean "representationalism". To effectively rebut Hansen's claims, however, we require an alternative explanation of those features of Putnam's dialectical practice that the ascription of the "Lockean" premiss purports to explain. Since the most salient feature of that practice is the apparently cavalier disregard of possible "non-linguistic" constraints on interpretation - the kinds of constraints proposed by exponents of the "canonical" strategy - an answer to Hansen may also illuminate more general matters of dispute between Putnam and his "canonical" critics.

At the beginning of this chapter, I claimed that commentators on the "model-theoretic" argument have often failed to pay sufficient attention to the argument's dialectical context. In the following pages, I shall endeavour to provide some support for that claim. The purpose

of the "model-theoretic" argument, as we have seen, is to undermine
metaphysical realism while leaving internal realism unscathed. Hansen
claims that Putnam smuggles certain additional substantive assumptions
into his account of metaphysical realism - assumptions that are neither
part of nor implied by the central tenets of what Hansen terms 'minimal
realism'; and he states that he is only concerned to defend minimal
realism against Putnam. But the central tenets of minimal realism, as
identified by Hansen, seem to be identical to those of the metaphysic-
al realism characterised earlier in this chapter, namely, a commitment
to "the notion of a mind-independent world and to a non-epistemic notion
of truth."[53].I assume that Hansen's "minimal realism" is, like Putnam's
"metaphysical realism", to be understood as a model of what it is for
any theory to be correct. If "minimal realism" were intended as an em-
pirical theory of some sort, then - like internal realism - it would
not be Putnam's intended target, and Hansen's claim that the "model-
theoretic" argument doesn't threaten minimal realism would be true (or
at least granted by Putnam) but inconsequential. I further assume that
"mind-independence", for the minimal realist, consists in being inde-
pendent of any representation that we may construct. On the "Lockean"
reading of Putnam, of course, being "mind-independent" will simply be
a matter of not being a mental state, since "representations" are to
be restricted to psychological states in the "narrow" sense.

Let us retain Putnam's term THE WORLD to stand for whatever it is
that serves as the realist's "mind-independent" reality. The realist's
contention is that truth is a matter of correspondence to THE WORLD,
and correspondence depends upon a determinate relation of reference
between terms and entities in THE WORLD. Putnam's challenge to the

realist is to explain how such a determinate relation of reference is produced. But, and this point is crucial, the realist requires that the constraints on reference allow for the possibility, at least in principle, that, given the extensions assigned to the terms in our language, even an "Ideal" theory in our language could be false of THE WORLD. This point is not undermined by Hansen's claim that the realist is not committed to a general fallibilism concerning our beliefs about "non-representational reality". For, as I argued above in discussing Devitt (section V-c), the realist's concession that some extension of our present methods of justification might be sufficent to determine the truth of some (or even all) such beliefs does not bridge the conceptual gap between realist truth and warranted assertibility. The realist remains committed to the possible falsity of an "Ideal" theory, and must therefore offer an account of the constraints on reference that is compatible with such a possibility. To put this point another way, the realist's purposes will be served only if he can show how we can determinately refer to classes of entities concerning whose members even our most warranted beliefs could at least in principle be false: this is what the "mind-independence" of THE WORLD amounts to.

Consider, now, as an example, Hacking's account, cited above, of the manner in which the realist is able to secure determinate reference. The necessary constraints are provided, according to Hacking, by the ways in which "language is embedded in a wide range of doings in the world", and he offers, as examples of such "doings", the fact that we eat the things we call cherries and stroke the things we call cats. The salient question, however, is whether the constraints cited by

Hacking suffice to determine referents that are "mind-independent" in the sense just defined. The suggestion seems to be that the extension of our term 'cat', for example, is composed of just those things with which we interact, or could interact, in the sorts of ways specified by Hacking. The class of cats is something that we carve out of "noumenal dough" through our practice. But we might ask whether it can sensibly be hypothesised that some of the things we _treat_ as cats - things with which we engage in standard human-feline interactions - might _really_ not be cats at all. Perhaps, to borrow from Putnam's discussion of a different problem (in the "pre-critical" paper, "Is Semantics Possible?"), some of these feline-like entities are really robotic Martian spies. Clearly, a view such as Hacking's could accomodate such a possibility by the simple expedient of enlarging the class of relevant human-feline interactions to include not only such benign "doings" as stroking and feeding, but also the more devious "doings" of actual or possible laboratory scientists. But suppose we further specify that the non-cathood of some of the things determined to be cats by actual or possible practice be in principle, _undetectable_ to us - that even an "idealisation" of our interactive practice would not lead us to doubt the cathood of these actual non-cats. _This_ possibility is surely not one that the view under consideration could accomodate. But it is precisely this sort of possibility that _must_ be accomodated if the class of cats singled out by Hacking's proposed constraints is to be "mind-independent" in the required realist sense.

It might be thought that Hacking can preserve the "mind-independence of the class of cats picked out by his constraints by maintaining that, even if our most warranted beliefs cannot mis-specify the member-

ship of that class, they <u>can</u> falsely ascribe certain properties to the members. We may include, in our "Ideal Cat-Theory", the assertion that all cats have a property ∅ when at least some cats <u>fail</u> "in reality" to possess such a property, it might be claimed. But it should be readily apparent that the argument we have just run for the class of cats can equally well be run for the class of things having the property ∅ If the constraint upon determinate reference is provided solely by "non-linguistic practice" of the sort cited by Hacking, the class of things having the property ∅ will be precisely the class of things we would ideally carve out via the relevant fragment of our practice, and our assertion, within our "Ideal Cat-Theory", that all cats have property ∅ will be <u>true</u>. What this shows is that an appeal to "non-linguistic facts" is insufficient to provide the sort of determinate reference-relation that the realist requires; because the reference-classes picked out by such facts are not "mind-independent" in the sense distinguished above. What the realist wants to say, of course, is that, in using our language in a given manner as part of our ongoing interaction with experienced objects, we bring it about that terms in our language "hook onto" classes of things that <u>are</u> "mind-independent" in the required sense - classes whose membership, and the properties of whose members, may, at least in principle, be incorrectly represented in even our most warranted theories. But this will require recourse to additional constraints - to classes in THE WORLD that pre-exist our employment of representational schemata in our attempts to carve the "noumenal dough", and to "causal relations" between <u>these</u> classes and terms in our language. Before examining the plausibility of an appeal to such additional constraints, however, let me return to

Hansen's criticisms of Putnam.

We are now, I think, in a position to explain those features of the "model-theoretic" argument that Hansen finds puzzling. He asks why Putnam ignores the possibility that the realist can provide the necessary additional constraints on reference by appealing to the sorts of "non-linguistic facts" cited by Putnam himself in his critique of "descriptivist" theories of reference. As we have just seen, however, "non-linguistic facts" about the embedding of language use in the broader framework of human activity cannot, by themselves, provide the kind of constraint that the metaphysical realist requires. But, we may now note, they do provide precisely the sorts of constraints that might be incorporated in an internal realist account of how terms acquire a determinate reference:

> For an internalist..., signs do not intrinsically correspond to objects, independently of how those signs are employed and by whom. But a sign that is actually employed in a particular way by a particular community of users can correspond to particular objects within the conceptual scheme of those users. "Objects" do not exist independently of conceptual schemes. We cut the world up into objects when we introduce one or another scheme of description. Since the objects and the signs are alike internal to the scheme of description, it is possible to say what matches what. 54

Indeed, internal realism can also incorporate the sort of "causal" account of how representations are attached to the world that writers like Devitt and Glymour have urged as counters to the "model-theoretic" argument. For, we may recall, internal realism as defined in "Realism and Reason", preserves the empirical content of "sophisticated realism", and the latter includes precisely the sort of scientific picture of linguistic representation urged by Putnam's realist opponents. Immed-

iately after rejecting an appeal to the "causal theory" of reference
as an answer to the "Models and Reality" presentation of the "model-
theoretic" argument, for example, Putnam volunteers the following obser-
vations:

> This is not to say that the construction of such a theory
> would be worthless as philosophy or as natural science.
> The program of cognitive psychology already alluded to,
> the program of describing our brains as computers which
> construct an "internal representation of the environ-
> ment", seems to require that mentalese utterances be,
> in some cases at least, describable as the causal product
> of devices in the brain and nervous system which "trans-
> duce" information from the environment, and such a des-
> cription might well be what the causal theorists are
> looking for. And the program of realism in the philo-
> sophy of science - of _empirical_ realism, not metaphys-
> ical realism - is to show that scientific theories can
> be regarded as better and better representations of
> an objective world with which we are interacting; and
> if such a view is to be part of science itself, as
> empirical realists contend it should be, then the inter-
> actions with the world by means of which this represen-
> tation is formed must themselves be part of the subject
> matter of the representation. But the problem as to how
> the _whole representation_, including the empirical theory
> of knowledge that is a part of it, can determinately
> refer is not a problem that can be solved by developing
> more and better empirical theory.   55

These remarks may be instructively compared with Clark Glymour's
characterisation of the distinctive beliefs of _metaphysical realists_:

> [Metaphysical realists] believe that there is a world
> of things and their properties and relations and that hu-
> man beings are a part of that world, a part which en-
> ters into special relations with itself and with some
> of the remainder. Humans may refer to parts of the world,
> describe the world, be acquainted with the world, and
> our sciences aspire to discover not only what the in-
> human features of the world may be, but also the struc-
> ture and constitution of the special relations between
> the world and those creatures that are sentient and
> articulate. "Metaphysical realism" is no more than the
> framework presupposed by every kind of scientific
> realism.   56

3

1.0	**2.8**	**2.5**
	3.2	**2.2**
	3.6	**2.0**
1.1	**4.0**	
		1.8
1.25	**1.4**	**1.6**
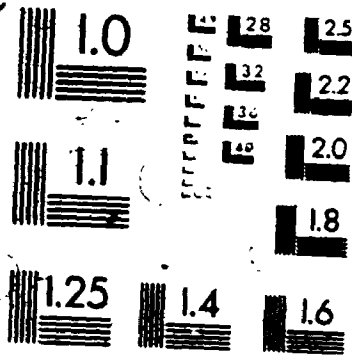
MicroD

Glymour's claim - that scientific realism presupposes metaphysical realism - is precisely what Putnam wishes to deny in "Realism and Reason". Not that the scientific realist should <u>reject</u> the framework of beliefs outlined by Glymour. But, so Putnam maintains, the scientific realist should construe talk of 'the world' and of our interactions with its inhabitants "internally", rather than in terms of THE WORLD of the metaphysical realist.

Far from ignoring the reference-determining potential of "non-linguistic facts", or tacitly recanting his earlier critique of "descriptivism", therefore, Putnam takes such things to be elements in the sort of "internalist" picture that he endorses. He does not regard such resources as being of any assistance to the cause of <u>metaphysical</u> realism, however, because they cannot determine reference-classes that are "mind-independent" in the required sense. It is because the realist requires reference to classes in THE WORLD that he needs to avail him- self of constraints other than our causal, or quasi-causal, interactions with things in the course of our everyday and scientific practice. And, Putnam maintains, the traditional realist solution to this problem has been to appeal to "operational and theoretical constraints". The "Skol- emisation" argument of "Models and Reality" purports to show that such an appeal fails to establish determinate reference to classes in THE WORLD, <u>whether these classes are taken to consist of theoretical entities, ordinary-sized physical objects, or past and future sense- data</u>. The problem, then, is <u>not</u> the "Lockean" one of determinately referring to extra-mental entities given that we have direct access only to our own ideas: the problem, rather, is to secure determinate reference to <u>anything that satisfies the realist's notion of "mind-</u>

independence", anything in THE WORLD. One way of setting up the problem

is indeed in terms of a "Lockean" epistemological premiss: assume that

we directly interact only with our own inner states, and try to estab-

lish the possibility of determinately referring to external objects

that are "mind-independent" in the realist's sense. But we can equally

well set up the problem with a "Kantian veil" rather than a "Lockean"

one: assume that we directly interact with external objects in sense

experience, and try to establish the possibility of determinately

referring to noumenal objects. What Hansen misses, in his Lockean-

sounding quotations from Putnam's discussion of the "model-theoretic"

argument, are Putnam's clarificatory remarks as to the sense to be

given to talk of "external objects". Thus the following passage is cited

in support of the "Lockean" interpretation:

> To pick out just one correspondence between words or men-
> tal signs and mind-independent things we would have
> already to have referential access to the mind-indepen-
> dent things. You can't single out a correspondence
> between two things by just squeezing one of them hard
> (or doing anything to just one of them)...   57

Hansen omits the remarks that immediately follow this passage, however:

> ...You cannot single out a correspondence between our
> concepts and the supposed noumenal objects without
> access to the noumenal objects.   (Stress added)

V-e   Appealing to THE WORLD

The challenge that the "model-theoretic" argument presents to the externalist is to furnish some account of the kinds of constraints that are to produce determinate reference relations between the terms in our language and "mind-independent" entities in THE WORLD. "Operational and theoretical constraints" are inadequate to the task, Putnam argues. And, as I argued in the previous section, an appeal to "non-linguistic facts" about the ways in which language use is embedded in our more general commerce with the world - whether these "facts" be features of our everyday activities or features of our scientific practice - fails to deliver the sorts of reference-classes that the realist requires. At this point in the debate, realists have tended to fall back on a further kind of constraint - one that is imposed by THE WORLD itself. As we noted in chapter 3, Putnam and his realist critics have engaged in a dialectical dance with the following steps. In response to the claim that the necessary constraints on reference are given by a "causal theory" of reference, or to Field's contention that reference is a "physicalistic" relation that we can define in terms of some physical parameter R, Putnam maintains that such constraints are "just more theory" that can be incorporated into our "ideal" theory to form an augmented theory for which we can still find appropriate SAT interpretations. The realist counters that, for a given proposed constraint C, "the constraint is not that an intended interpretation must somehow make our account of C come out true. The constraint is that an intended interpretation must conform to C itself."[58] In terms of Field's thesis, the counter is that it is not the requirement that the definition of 'reference' be satis-

fied that yields a determinate interpretation of the language; rather, it is R itself, as a relation in THE WORLD, that determines such an interpretation, including the interpretation of the terms used in giving the physicalistic definition of 'reference'.

Putnam responds to this claim by charging that the realist is "ignoring his own epistemological position"[59] by assuming that his own use of the term 'reference', or of the terms employed in formulating 'C' or 'R', determinately and uniquely refers. For there would seem to be as many different "reference" relations as there are SAT interpretations of IT*. In each SAT model of IT*, M_i, the term 'reference' in our language will be associated with a relation R_i. Furthermore, if we can make sense of the idea that something can stand in a given relation to that relation itself, then, in each M_i, 'reference' will stand in R_i to the relation R_i itself.[60] If the realist's use of the term 'reference' (or of the terms used in the formulation of 'C' or 'R') determinately refers to just one of these R_i, say R_{17}, then 'reference' must be attached to R_{17} by "metaphysical glue", Putnam maintains. To assume that one's own terms determinately refer is tantamount to begging the very point at issue, namely, how any terms can determinately refer in the manner required by the realist.

The realist responds by returning the charge of question-begging. Lewis, for example, suggests that it is unfair of Putnam to require that the realist (whom he terms "the Respondent") give an account of the manner whereby the terms employed in giving the necessary "constraint" on interpretation themselves determinately refer, where this account is not allowed to appeal to the very constraint formulated by means of these terms:

> The respondent may indeed have given a correct account
> of the constraint that makes determinate reference poss-
> ible, couched in language that does indeed have determin-
> ate reference in virtue of the very constraint that it
> describes. 61

Devitt is characteristically more forthright in making a related point:

> Putnam claims to be offering an argument against "meta-
> physical realism". At no point is he entitled to <u>assume</u>
> this doctrine false. If the doctrine is true, then there
> will be determinate referential relations between the
> words of any theory and pieces of the world. This will
> be true also of the words of any theory of reference
> used to explain these relations. If such a theory is
> comprehensive it will of course apply to its own words.
> Putnam's anti-realist argument depends on there being
> no answer to the question about what determines ref-
> erence for [the 'Ideal" theory]. Using a theory of ref-
> erence there is an answer: reference is determined by
> causal relations of a certain sort. That answer works
> for 'causally related' just as it does for 'cat'. Putnam
> begs the question by simply assuming 'causally related'
> lacks determinate reference. 62

What fixes the reference of the term 'reference', according to
Devitt, is "causal glue", not "metaphysical glue". Putnam is unconvin-
ced, however, maintaining that Devitt's "causal glue" is just metaphys-
ical glue bearing a less contentious label. If it is the case that every
R_i generated by a particular SAT interpretation of IT^* is in its own
converse domain if the realist's chosen R_{17} is, then it cannot be this
formal fact about R_{17} that picks it out as <u>the</u> relation of reference.
Rather, the realist has to hold that it is simply what Putnam terms a
"surd, metaphysical fact" = an ultimate fact about THE WORLD - that R_{17},
rather than any other R_i, just <u>is</u> reference. If R_{17} <u>is</u> the referent of
our term 'reference', and if R_{17} is indeed a causal relation, then it
will indeed be the case that the term 'reference' is tied to R_{17} by

"causal glue". But the satisfaction of the first antecedent requires a prior metaphysical fact about the way in which language acquires its interpretation, and it is just the appeal to such a fact that Putnam has in mind when he talks about a need for "metaphysical glue". He regards the idea that a sign-relation is built into THE WORLD itself as a return to a form of "medieval essentialism", where one holds that there are "self-identifying objects" and that THE WORLD itself "sorts things into kinds."[64] He is thus no more sympathetic to the Merrill-Lewis line. sketched in the previous section, the line that takes the realist's saving constraint to be provided by the "built-in structure" of THE WORLD. Talk of "eligible referents" and objective "joints" in THE WORLD is dismissed as "spooky" and "medieval-sounding".[65]

Let us take stock of this exchange between Putnam and his realist critics. Suppose that we grant Putnam his contention that the realist is committed to the existence of "surd, metaphysical facts" concerning the semantic properties of language, and to "spooky" and "medieval-sounding" theses relating to the existence of objective "joints" in THE WORLD. Why should any of this worry the metaphysical realist? (Lewis's response to the charges of "spookiness" et.al. is: "Well, sticks and stones may break my bones..."!) Isn't the realist perfectly happy to entertain the possibility of "surd metaphysical facts", or of objective "joints" in THE WORLD? Putnam's charge, in "Realism and Reason", is that the "model-theoretic" argument demonstrates the incoherence of metaphysical realism, but it now begins to look as if the latter position is perfectly coherent just so long as the realist is prepared to endorse exactly the sorts of claims that might be taken to be definitive of metaphysical realism in the first place! Does Putnam give us any further

reason to question the coherence of a metaphysical realism that, quite consistently it would seem, answers Putnam's accusations in this way? Before examining certain supplementary arguments that bear upon this question, I want to assess the independent viability of two distinct realist conceptions of what it is in THE WORLD that provides the needed constraints on interpretation - the conceptions attributed, in my earlier sketch of the "canonical" strategy, to Devitt and Glymour, on the one hand, and to Merrill and Lewis, on the other.

Devitt, Glymour, and Field are proponents of what I shall term the "reference approach" - the view that it is reference itself, as a particular relation between language and THE WORLD, that provides the additional constraints on reference that the realist requires. Reference, so construed, is a relation that can - in practice (Field) or in principle (Glymour) - be defined in terms of the fundamental magnitudes of physical science. Merrill and Lewis, on the other hand, adhere to what I shall term the "structure approach" - the view that the necessary constraints on interpretation are provided by "structure in THE WORLD" - certain objective relations holding between entities in THE WORLD, and certain objectively correct ways of sorting such entities into classes. We might enquire whether either of these approaches, in isolation, is adequate to the realist's purposes.

That the "reference" approach requires supplementation by the resources of the "structure" approach is readily apparent if, following Putnam, we reflect upon the fact that the extensions of even such simple terms as 'horse' or 'rabbit' include

many things we have not causally interacted with (e.g.

> future horses and rabbits, or horses and rabbits that
> never interacted with any human being). When we use the
> word 'horse' we refer not only to the horses we have a
> real connection to, but also to all other things of the
> same kind...The "metaphysical realist" solution (to the
> problem of how this is possible)...is to say that the
> word automatically covers not just the objects I lassoo-
> ed (by causal interaction), but also the objects which are
> of the same kind - of the same kind in themselves. But
> then the world is...being claimed to contain Self-Iden-
> tifying Objects, for this is what it means to say that
> the world, and not thinkers, sorts things into kinds.   66

In other words, the "reference" approach, by itself, will not secure

for the realist the sorts of reference-classes in THE WORLD that he

requires, even if it secures reference to entities in THE WORLD.

Equally inadequate, I think, is the "structure" approach unbul-

warked by the "reference" approach. For, even if there is "structure

in THE WORLD, there remains the question how it comes about that all

admissible interpretations of our language respect such structure.

Both Merrill and Lewis think this can be done by appeal to the exist-

ence of formal relations between structures, even if such relations

elude our recognitional capacities. All the realist requires, accor-

ding to Merrill, is the possibility that, given an "Ideal" theory,

there might be no isomorphic mapping of the language of that theory

onto THE WORLD that made the theory come out true. Nothing in this

account, he claims, requires that we be able to pick out, by our use

of the language or by our standing in a particular causal relation to

THE WORLD, a unique interpretation of the language of the theory. It

is quite acceptable to the realist, Merrill maintains, that there

might be more than one true interpretation of a theory satisfying the

realist constraint. There is no onus upon the realist, in such a case,

to show that only one of these interpretations is <u>really</u> correct.[67]
Lewis similarly allows that there may be "modest indeterminacy of
reference" given the class of "eligible interpretations", where the
latter is composed of interpretations that maximise the satisfaction
of the twin desiderata that the referents be eligible and the theory
be true.[68]

The problem with this strategy, however, as Gregory Currie has
pointed out, is that it presupposes that the interpretation of our
<u>semantical</u> terms can be fixed in a way that circumvents Putnam's
argument. For, unless the interpretation of the terms that Merrill
employs in formulating his constraint has <u>already been established</u>,
we can simply run Putnam's "just more theory" response: given that
our "Ideal" theory conjoined with the proposed semantic constraints
provided by the "structured domain" semantics is consistent, we can
find a SAT interpretation of the augmented theory IT* such that the
"Ideal" theory is true and the semantic constraints (as interpreted)
are satisfied.[69] Similar problems beset Lewis's proposal, in spite of
his adamant protestations to the contrary:

> The proposed constraint is that referents are to be
> eligible, not just that eligibility theory is to be
> satisfied somehow, not just that the referents of
> 'cat', etc., are to be included among the referents
> of 'eligible'.   70

For, unless we are already given a determinate interpretation of the
term 'eligible', the latter is just as much 'up for grabs', semantically
speaking, as any other term in the language, and the "just more theory"
response can prevail.

The obvious recourse for the proponent of the "structure" approach is to draw upon the resources of the "reference" approach, but the distinctive features of the former, as expounded by Merrill and Lewis, fail to survive such a marriage of convenience. For if there just are, as a "surd metaphysical fact", reference relations between particular terms in our language and particular entities in THE WORLD, then that M is an isomorphic mapping from the language of IT onto THE WORLD is no longer sufficient to make M an acceptable interpretation of IT, for M may not respect the "surd metaphysical facts" about reference. In fact, the only sort of account that promises to be adequate to the realist's purposes seems to be the kind of hybrid sketched by Putnam in his argument against the independent viability of the "reference" approach, cited above. The realist must hold that reference is a physical relation between terms in our language and particular entities in THE WORLD, and that the extensions of those terms transcend those particular entities in virtue of (a) the existence of objective "joints" in THE WORLD separating things into kinds, and (b) our intention that our terms should pick out such kinds in THE WORLD. The question, then, is whether such a realist account is both coherent and plausible.

In chapter three above, I remarked an argument, in the first of the Howison Lectures, to the effect that such an account is not coherent. The incoherence purportedly resides in the attempt to construe reference, or the sort of causal relation in terms of which writers such as Devitt and Field propose to elucidate reference, as a physical relation in THE WORLD, where the latter is taken to be describable, at least in principle, purely in terms of a materialist ontology. Let me briefly summarise the features of Putnam's argument that bear upon our

present concerns. He begins by drawing a distinction between two senses in which we talk of one thing being the 'cause' of another. On the one hand, there is what Mill termed a "total cause", where to say that A caused B is to say that the occurrence of an A-type event is a <u>suffic-icient</u> condition for the occurrence of a B-type event. But we rarely use 'cause' in this sense. Rather, "we regard certain parts of the total cause as 'background', and refer only to the part of interest as 'the cause'."[71] To say that A caused B in this second sense of 'cause' is tantamount to saying that we can <u>explain</u> B in terms of A, and it is this sense of 'cause' that features in "causal theories" of reference:

> When it is said that a word refers to x just in case the (use of the) word is connected to x by a "causal chain of the appropriate type", the notion of a "caus-al chain" involved is that of an <u>explanatory</u> chain. Even if the notion of "total cause" <u>were</u> physically definable, it would not be possible to <u>use</u> it either in daily life or in philosophy; the notion the mater-ialist really uses when he employs 'causal chain', etc., in his philosophical explications is the intui-tive notion of an <u>explanation</u>. [72]

Putnam argues that the "explanatory" notion of cause cannot be elucidated in physicalist terms. It cannot be <u>defined</u> in such terms for two reasons. Firstly, the notion is "abstract" in the sense that we could apply it even in a possible world containing non-physical things or properties. Secondly, as Putnam argued in the <u>Locke Lectures</u>, what we count as an acceptable explanation depends upon our interests and background knowledge, and thus upon the relative projectibility of different predicates. Given Goodman's argument that no purely <u>formal</u> criterion can serve to distinguish between projectible and non-projec-

tible predicates[73], we cannot evade the first difficulty by character-
ising the "abstract" notion of cause in equally abstract terms, making
use of a purely formal criterion. Putnam further maintains that the
materialist cannot take causation, in the "explanatory" sense, as a
primitive notion, for this would be incompatible with the materialist's
commitment to physicalism:

> "Causal powers" are properties that explain something,
> given background conditions and given standards of sal-
> ience and relevance...Salience and relevance are attrib-
> utes of thought and reasoning, not of nature. To pro-
> ject them into the realist's "real world" into what
> Kant called the noumenal world, is to mix objective
> idealism (or perhaps medieval Aristoteleanism) and mat-
> erialism in a totally incoherent way...If events in-
> trinsically explain other events, if there are saliences,
> relevances, standards of what are "normal" conditions,
> and so on, built into the world itself independently
> of minds, then the world is in many ways like a mind,
> or infused with something very much like reason. And if
> that is true, then materialism cannot be true.   74

A possible realist rejoinder to the preceding argument leads Putnam
to consider the materialist view of reference. The linking passage,
which requires some elucidation, runs as follows:

> Some metaphysical materialists might respond to what has
> been said by agreeing that "A causes B" does not describe
> a simple relation between A and B. "All you are saying
> is that causal statements rest on a distinction between
> background conditions and differentiating factors, and I
> agree that this distinction isn't built into the things
> themselves, but is a reflection of the way we think
> about the things," such a philosopher might say. But
> here he has used the words 'think about', i.e. he has
> appealed to the notion of reference.   75

Putnam's suggestion, if I understand him correctly, is that the
materialist might concede that causation, in the "explanatory" sense,
cannot be a physical relation, while continuing to maintain that there

is a "built-in structure" to THE WORLD, a structure that includes "total causes". In advancing such a thesis, the materialist avails himself of the distinction between properties that THE WORLD has "in itself" and properties that are merely functions of "the way we think about THE WORLD", thereby reflecting our interests and purposes. Putnam's point, then, is that the materialist is entitled to such a distinction only if he can provide an account of reference that circumvents the problems raised by the "model-theoretic" argument. And the ensuing argument, not surprisingly, is to the general effect that all the difficulties seen to beset the attempt to construe cause, in the "explanatory" sense, as a physical relation also beset the attempt to construe reference in this fashion. Reference, Putnam argues, is just as "abstract" a notion as causation, and "a definition of reference from which it followed that we could not refer to a non-physical magnitude if there were one is just wrong."[76] Nor can we take reference as a physical primitive, for, like causation, it is

> a flexible, interest-relative notion: what we count as
> referring to something depends on background knowledge
> and our willingness to be charitable in interpretation.
> To read a relation so deeply human and so pervasively
> intentional into the world and to call the resulting
> metaphysical picture satisfactory (never mind whether
> it's "materialist") is absurd.   [77]

If reference is indeed as "deeply human and pervasively intentional" as Putnam claims, then an attempt to preserve externalism by recourse to constraints on interpretation furnished by reference construed as a physical relation in THE WORLD would, as Putnam says, be incoherent. But must the realist accept these claims about reference? Putnam bases the latter upon what he terms "the flexible, interest-

relative nature of reference, exemplified by the fact that "what we count as referring to something depends on background knowledge and our willingness to be charitable in interpretation" (added stress). But these are features of our practice of assigning referents to terms employed by other speakers of our own language or by speakers of other languages - earlier generations of scientists, members of alien cultures, or hypothetical automata whose behaviour mirrors our own. Why does Putnam believe that these "interest-relative" features of our reference-ascribing practices render reference itself "interest-relative"? The answer, I think, is that Putnam is implicitly drawing on the account of translation - of assigning referents and "meanings" to terms and expressions in a foreign language - presented in the Locke Lectures. Putnam's contention, we may recall, is as follows:

> I am not just contending that it is good methodology
> in finding out what a speaker "means" to try to rationalize
> his behaviour in this way! I am suggesting that what it
> is to be a correct translation or truth-definition is
> to be the translation or truth-definition that best
> explains the behaviour of the speaker. This is a sub-
> stantive metaphysical theory of what "correctness" is
> in linguistics...Now, rationalizing the behaviour of
> a speaker is explaining...Explaining a large number of
> facts about a speaker's behaviour-dispositions...is,
> if I mean by the big pretentious expression 'ration-
> alize behaviour'. But explanation is an interest-relat-
> ive notion.

The claim is that the correct reference assignment for an alien term just is the assignment that provides, together with auxiliary assumptions, the best "rationalization" of the speaker's behaviour: there is no "fact of the matter" about reference independently of our reference-ascribing practice. And, for familiar Quinian reasons[79], what holds for translation from a foreign language also holds for other speakers of

our own language, where we apply the principle of "homophonic trans-
lation". But, Putnam maintains, our practice of rationalizing the
behaviour of speakers is _explanatory_ in its intentions, and explanation
is interest-relative. Thus we derive the required conclusion that _refer-
ence itself_, as a function of our reference-ascribing practices, is
interest-relative.

But why should the externalist accept Putnam's "substantive meta-
physical theory" about the nature of "correctness" in linguistics?
Suppose we grant that our practice of ascribing referents to the terms
employed by linguistic agents is part of our broader attempts to make
sense of behaviour as a means to effectively interacting with one
another. The importance of reference-ascription, so construed, is
surely that it allows us to see others as trying to cope with the same
world that we ourselves inhabit - as referring, by means of their lan-
guage, to the same things that we refer to by means of ours. Suppose
we also grant that the only available standards for _judging_ our refer-
ence-ascriptions are those encompassed by the broader context of attemp-
ted rationalization. Still, the externalist will maintain, this is
quite consistent with the further contention that there is an independ-
ent "fact of the matter" as to whether those reference-ascriptions
that satisfy such standards _really_ get it right - where "getting it
right" is a matter of matching the term in _their_ language that stands
in the physical relation of reference to a given entity in THE WORLD
with the term in _our_ language that stands in the same relation to the
same entity. In identifying the interest-relativity inherent in our
reference-ascribing practices, Putnam may indeed have revealed our

epistemic limitations in respect of providing an account of how our languages map onto THE WORLD, the externalist may claim, but this in no way demonstrates the incoherence of the proposed externalist account of reference itself.

Putnam's other argument against a physicalistic account fares even worse, I think. He contends that such an account cannot do justice to the "abstract" nature of reference, where the latter is taken to involve the possibility of referring to "non-physical things and properties", were such to exist. But there seems to be a confusion, here, between (a) the question whether reference can be a physicalistic relation, and (b) the question whether it is possible to refer to non-physical things. A positive answer to the first question will entail a negative answer to the second only if it is also held that non-physical things cannot enter into physicalistic relations with other things. But the denial of the latter thesis, while it may generate difficulties (witness the traditional problems associated with Cartesian dualism), is surely not incoherent. Indeed, if it is to be possible for us to interact in some way with "non-physical things or properties", through the mediation of their manifest effects, then it seems that it must be possible for them to enter into some physicalistic relations in order to bring about those effects (or our experience of those effects). But, in that case, it is not obvious that there is a special problem with the idea that reference is a physicalistic relation that can obtain between terms in a language and non-physical entities. Reference, as a physicalistic relation, might be taken to be parasitic, in such cases, upon the other physicalistic relations whereby non-physical things are able to bring about physical

effects. This, of course, will not explain how we could refer to non-physical entities that do _not_ enter into the latter kinds of physicalistic relations, But it is unclear how we could interact in any way with non-physical entities of _this_ sort. And, if we cannot interact with them at all, then the externalist can appeal to Putnam's own reasoning concerning the referential capacities of envatted brains to show that we _couldn't_ refer to such non-physical things. Thus the possibility of referring to the non-physical doesn't seem to present any special problem for a physicalistic construal of reference.

V-f   Conclusions

In evaluating the critical debate over the merits of the "model-theoretic" argument, and in examining (in V-a, sub-section 2, above) certain auxiliary anti-realist arguments, I have suggested that the externalist can escape the charge of incoherence by arguing that the anti-realist arguments inflict only "epistemological damage" on realism and leave the metaphysics of realism unscathed.[80] There is, it seems, nothing incoherent in the claim that THE WORLD has a determinate "built-in" structure; that it comprises determinate classes of entities that stand in the physical relation of reference to linguistic and other representations employed by symbol-using creatures; that it is capable of rendering false even an "Ideal" theory, and of rendering correct only one of a number of "empirically equivalent" theories. All that Putnam's arguments demonstrate, the realist may claim, is that our ignorance of THE WORLD is perhaps much more profound, and our practices of forming beliefs, selecting theories, and ascribing referents, permeated by our interests to a much greater extent, than realists have hitherto supposed. But this poses no conceptual problem for the realist, who is already committed to the existence of a gap between epistemic warrant and truth. We have also seen that the externalist can respond to the charge that the existence of such a gap undermines the expressibility of the realist thesis by maintaining that such a charge is question-begging, for it rests upon the assumption that the term 'reference', as employed by the externalist, cannot determinately refer to a specific relation in THE WORLD via the mediation of precisely that relation. The epistemological consequences of externalism, the realist may conclude, in no way entail either the

incoherence or the falsity of that doctrine.

Can the externalist, then, simply ignore the epistemological im-
plications of realism? If we focus not upon "Realism and Reason", where
the charge of "incoherence" is levelled, but upon RTH, I think we find,
at least implicitly, the suggestion that the epistemological consequen-
ces of realism are significant not because they demonstrate its inconer-
ence, but because they point to its philosophical bankruptcy. So viewed,
the "model-theoretic" argument and the auxiliary arguments examined ear-
lier take their place alongside the discussion of the skeptical puzzles
generated by externalism (see III-b, sub-section 3, above) as evidence
in the case against one research programme in philosophy - the external-
ist programme in metaphysics and the philosophy of language - and, in-
directly, in favour of an alternative research programme - one generated
by the adoption of the internalist perspective - for which such exter-
nalist "anomalies" do not arise.

Suppose we use 'R' as a label for the realist's physicalistic rel-
ation between terms in a language and entities, or classes of entities
in THE WORLD, and suppose we use the label 'R-Truth' to stand for the
satisfaction-relation that may exist between sentences in L and, states
of THE WORLD under the mapping provided by R. The realist contends
that there is nothing inconerent in the supposition that such a relat-
ion R exists, and that even an "Ideal" theory might fail to be R-true.
But, as we have seen, the epistemological consequences of externalism
include our inability to determine either (a) whether terms in differ-
ent languages, or different uses of a term in a single language, stand
in R to the same entity in THE WORLD, and thus have the same reference;

or (b) which, if any, of our beliefs about the world - that is, THE
WORLD - are true - that is, R-True; or (c) which of a set of "empir-
ically equivalent" theories is really-true (R-True). Given these con-
siderations, however, we might ask whether the proposal that we construe
reference as R and truth as R-Truth offers any promise of illuminating
the sorts of questions that lead us to seek a philosophical account of
reference and truth in the first place. The latter, it would seem, are
questions that pertain to the nature of linguistic understanding and
use, and to the place of linguistic practice in the broader context
of human activity. Can we expect the realist account of reference and
truth to throw any light on these matters?

We may distinguish two ways in which philosophers have attempted
to incorporate the notions of R and R-Truth into an account of linguis-
tic practice. First, there are attempts to vindicate these notions
within the framework of a traditional truth-conditional semantics. In
such a semantics, understanding a term in a language is a matter of
grasping its referent, or grasping the manner whereby its referent is
singled out, and understanding a sentence is a matter of grasping its
truth-conditions. The externalist takes the latter to involve a grasp
of the conditions under which a sentence is R-True. Since an account
of linguistic understanding is part of an explanation of speaker's
linguistic and non-linguistic behaviour, the notions of R and R-Truth
will play a central role in a philosophical account of language. But
does it make sense to attribute to speakers a capacity to grasp the
R-Truth conditions of sentences, given the radically non-epistemic
nature of R-Truth? In the following chapter, I shall examine in detail
Dummett's arguments for a negative answer to this question.

Secondly, one might try to make a place for the notions of R and R-Truth within a philosophical account of language by adopting Putnam's own strategy in "Reference and Understanding". It might be argued that, although we cannot take a speaker's understanding of a language to consist in a grasp of R-Truth conditions, the sentences of the language nonetheless possess such truth-conditions, and this fact is crucial to an understanding of the ways in which linguistic practice contributes to the success of our linguistically-mediated endeavours. I shall examine a variant of this strategy in the following chapter. We may note, however, that our examination of the "canonical" strategy in the previous two sections does much to undermine the appeal of this sort of externalist strategy, since, as we have seen, the "explanatory" virtues of the externalist account seem to be preserved - but without the epistemological costs - by internal realism.

If neither a truth-conditional semantics nor a philosophical account of the function and utility of language use proves to be a means of providing a respectable philosophical role for the notions of R and R-Truth, it seems open to us, in the absence of arguments to the contrary, to maintain that these externalist notions, even if coherent, fail to elucidate the notions of reference and truth that interest us as philosophers. This would be to side with Goodman and the earlier Rorty in holding that the externalist's WORLD, together with the notions of R and R-Truth that are parasitical upon it, is something that is indeed "well lost".[82] It is just such a conclusion that I shall urge at the end of the following chapter.

Endnotes - Chapter Five

1. Putnam, 1976d, pp. 131-132.

2. Ibid., p. 133.

3. See, e.g., Goodman, 1960.

4. Putnam, 1976d, p. 130.

5. Quine, 1969.

6. Putnam, 1976d, pp. 133-135.

7. See, for example, Melchert, 1986. For a critical discussion of Melchert's paper, see Davies, 1987a.

8. Putnam, 1980, p. 472.

9. Ibid., p. 477.

10. Ibid., pp. 481-482.

11. Glymour, 1982.

12. Lewis, 1984.

13. Putnam, 1980, p. 473.

14. Currie, 1982.

15. Smart, 1982.

16. Alan H. Goldman, 1984.

17. Smith, 1983.

18. Smart, 1982, p. 2.

19. See also my comments on Smith's response to Melchert, 1986 (in Smith, 1986) in Davies, 1987a.

20. Alan H. Goldman, 1984, p. 30.

21. Ibid., p. 28.

22. Ibid., p. 32.

23. Ibid., p. 33.

24. See especially footnote 2, p. 29, of Goldman's paper for this conflation. ...

25. Alan H. Goldman, 1984, p. 31.

26. Pearce and Rantala, 1982.

27. Hacking, 1983, pp. 104-105.

28. See especially Hansen, 1987, and Lewis, 1984. For an interesting discussion of historical antecedents of Putnam's argument, see Demopoulos and Friedman, 1985.

29. Hansen, 1987, p. 83.

30. Ibid., p. 84.

31. Devitt, 1983a.

32. Compare Crispin Wright's response (Wright, 1979) to a similar charge that Strawson levels at Dummett's arguments against semantic realism. This response is further discussed in sub-section 2, section VI-c, below.

33. Devitt, 1983a, p. 297.

34. Lewis, 1984, p. 224.

35. Ibid., p. 225.

36. Plantinga, 1982.

37. Alvin Goldman, 1986, p. 155.

38. Demopoulos, 1982, p. 138.

39. Ibid., pp. 138-139.

40. Hacking, 1983, p. 105. For a similar line, see Blackburn, 1984.

41. Glymour, 1972, p. 177.

42. Devitt, 1983a, p. 298. See also Brueckner, 1984.

43. Merrill, 1980, p. 72.

44. Ibid., pp. 73-74.

45. Lewis, 1984, p. 227.

46. See Putnam, 1975c.

47. Hansen, 1987.

48. Putnam, 1983a, p. xii.

49. Hansen, 1987, pp. 92-93.

50. Ibid., pp. 89-90.

51. Ibid., p 92.

52. Ibid , p. 93.

53. Ibid., p. 78.

54. Putnam, 1981a, p. 52.

55. Putnam, 1980, p. 479

56. Glymour, 1982, pp. 179-180.

57 Putnam, 1981a, p. 73, cited in Hansen, 1987, p. 92.

58. Lewis, 1984, p. 225.

59. Putnam, 1983a, p. xi.

60. See the discussion of Devitt in Putnam, 1983c.

61. Lewis, 1984, p. 226.

62. Devitt, 1983a, p. 299.

63. Putnam, 1983c, pp. 295-296.

64. Putnam, 1983a, p. xii.

65. Quoted by Lewis in Lewis, 1984, p. 229.

66. Putnam, 1981a, p. 53.

67. Merrill, 1980, pp. 74-76.

68. Lewis, 1984, pp. 227-228.

69. Currie, 1982.

70. Lewis, 1984, p. 228.

71. Putnam, 1982b, p. 213.

72. Loc. cit.

73. Goodman, 1954, chapter III.

74. Putnam, 1982b, pp. 215-216.

75. Ibid., p. 221.

76. Ibid., p. 223.

77. Ibid., p. 225.

78. Putnam, 1976a, p. 41.

79. See Quine, 1960, chapter 2.

80. See Lewis, 1984, p. 232, for a version of this strategy.

81. Putnam, 1976d, p. 137.

82. See Goodman, 1960, and Rorty, 1972.

Chapter Six   Dummett's "Anti-Realism"

VI-a   The Many Faces of (Dummettian) Realism

In section V-f above, I suggested that, while there may be nothing incoherent in the externalist contention that language stands in the relation R to entities in THE WORLD and that sentences thereby possess determinate R-Truth conditions and R-Truth-values, we might still be inclined to reject externalism on the grounds that it fails to illuminate any of the interesting philosophical questions about reference and truth. The salient question, I claimed, is whether the notions of R and R-Truth are usefully included in a philosophical account of the operations of language. The burden of the present chapter is to argue, following Dummett and Wright, that they are not usefully so included. As a preliminary to this, however, it is important to distinguish a number of "realism vs. anti-realism" debates that inter-twine, sometimes treacherously, in the writings of Dummett and his commentators.

Firstly, we have the familiar distinction between "metaphysical realism" and various forms of non-externalism. The issue here is whether truth is "radically non-epistemic". The realist, in this debate, identifies truth with R-Truth, whereas the anti-realist denies this identification. Secondly, we can distinguish two kinds of "realism" debates within the philosophy of language, where "realism", in each case, involves a particular application of the "metaphysical realist" notion of R-Truth. The "meaning-theoretic realist" affirms, and his anti-realist opponent denies, that understanding a language consists in a grasp of R-Truth conditions for the sentences of the language. The "semantic realist" affirms, and his anti-realist opponent denies, that an adequate philosophical account of language should associate R-Truth

210

conditions with the sentences of a language as (at least part of) an account of their "meaning" or "assertoric content". Finally, there are debates between proponents and opponents of various brands of "ontological realism", such as "scientific realism", "moral realism", and "mathematical realism". The issue, in such debates, is whether there "really are" - in some sense - entities of the kind hypothesised in a given mode of discourse.

Let me make a few observations concerning the relationships that obtain between these different kinds of "realism" and "anti-realism". Firstly, meaning-theoretic realism and semantic realism represent the two options canvassed at the end of the previous chapter whereby the externalist might establish the relevance of the notion of R-Truth to an adequate philosophical account of language. The strategy of rejecting meaning-theoretic realism but endorsing semantic realism is the one adopted by Putnam himself in "Reference and Understanding". Note that a non-externalist can adopt an analogue of this strategy, where, in place of the semantic realist's claim that sentences have content in virtue of relations obtaining between their components and THE WORLD, we have the claim that sentences have such content in virtue of relations with the world. Secondly, the reader will note that the characterisation of ontological realism makes no reference to R-Truth. One might wish to introduce the latter notion as a means to cashing out talk of there "really being" entities of a particular kind: one might say that there "really are" X's just in case the term 'X' stands in the relation R to some class of things in THE WORLD, or just in case a suitable subset of our beliefs about X's are R-True. But the characterisation of

ontological realism offered above leaves it open to one to <u>refrain</u> from explicating the doctrine in this way, and, indeed, to be an ontological realist with respect to a given class of statements without necessarily subscribing to <u>any</u> of the other forms of realism. Putnam, as we have seen, wishes to cleave to scientific realism (and, indeed, many other forms of ontological realism) while giving up metaphysical realism and thus all the other "realisms" that are parasitic thereupon. Crispin Wright is equally insistent that ontological realism does not entail metaphysical, meaning-theoretic, or semantic realism. He characterises the realist position, in debates over scientific realism, moral realism, etc., as follows:

> A realist about one of the areas will hold, as usual,
> that its statements are associated with determinate
> truth-conditions which the world may or may not objec-
> tively realize. If he is an anti-reductionist, he will
> hold in addition that no independent account can be given
> of these truth-conditions in terms of a more basic or
> epistemologically more favoured vocabulary; a realist
> of this sort about theoretical science, for instance,
> would hold not merely that our theories can be objec-
> tively true or false but also that the world "contains"
> states of affairs for whose satisfactory description
> the use of distinctively scientific-theoretical vocab-
> ulary is indispensable.   1

Realism so construed is quite compatible, Wright maintains; with the rejection of metaphysical realism and of a picture of language that em-ploys the notion of R-Truth; For the realist's talk of truth and fal-sity need not violate the (metaphysical) anti-realist's proscription on recognition-transcendent truth.

If Putnam and Wright observe the metaphysical neutrality of onto-logical realism, the same cannot be said for some of Dummett's critics[2], nor, perhaps, for Dummett himself, if we require that the neutrality be

clearly observed. I shall return to the failings of Dummett's critics
shortly. For the present, however, I shall attempt to clarify Dummett's
own position concerning the nature, plausibility, and interdependence
of the various forms of realism and anti-realism distinguished above.
When commentators speak of "Dummettian anti-realism", what they have
in mind (or _should_ have in mind) are a number of arguments, developed
most fully in the papers, "What is a Theory of Meaning? I" and "What
is a Theory of Meaning? II" (henceforth, WTM-I and WTM-II), that
challenge meaning-theoretic realism and, less explicitly, semantic
realism. I shall examine the details of these arguments in the following
section. We may, however, consider Dummett's general strategy and the
conclusions he claims to establish, in anticipation of the more detailed
treatment to come. Dummett follows Davidson, among others, in holding
that the solution to traditional problems in the philosophy of language
is to be sought through the construction of a "theory of meaning":

> The best method of formulating the philosophical prob-
> lems surrounding the concept of meaning and related
> notions is by asking what form should be taken by what is
> called a "theory of meaning" for any one entire language;
> that is, a detailed specification of the meanings of all
> the words and sentence-forming operators of the language,
> yielding a specification of the meaning of every express-
> ion and sentence of the language.   3

The goal is not the surely impractical one of actually _producing_ such
a "theory of meaning", but, rather, the enunciation of the general
principles that would be employed in the production of such a theory.
Dummett further maintains that the proper way to specify the "meanings"
of the elements in a language is to characterise what is involved in
knowing such meanings; and to know the meaning of a language is to
know how the language functions as a language, how it is _used_. This,

for Dummett, is at least part of what Wittgenstein meant by identify-
ing _meaning_ with _use_:

> A theory of meaning is a theory of understanding; that is,
> what a theory of meaning has to give an account of is
> what it is that someone knows when he knows the langu-
> age, that is, when he knows the meanings of the express-
> ions and sentences of the language....
> To grasp the meaning of an expression is to understand
> its role in the language: a complete theory of meaning
> for a language is, therefore, a complete theory of how
> the language functions as a language..., that is, of how
> its speakers communicate by means of it: here "commun-
> icate" has no more precise signification than "do what-
> ever may be done by the utterance of one or more sen-
> tences of the language." [4]

In "WTM-I", Dummett argues that a "theory of meaning" must be
what he terms "full-blooded" rather than "modest". He claims that,
if such a theory is to be a theory of understanding, then it "must,
in explaining what one must know in order to know the meaning of each
expression in the language, simultaneously explain what it is to have
the concepts expressible by means of that language." [5] A "modest" theory
of meaning is one that purports only to "give the interpretation of the
language to someone who already has the concepts required", whereas a
"full-blooded" theory also endeavours "to explain new concepts to some-
one who does not already have them." [6] Dummett argues that a modest
theory of meaning is functionally no different from a translation manu-
al, and therefore fails to provide the sort of philosophical illumin-
ation that we require from a theory of meaning. In "WTM-II" he further
argues that a full-blooded theory of meaning cannot be a form of
meaning-theoretic realism - that is, its central notion cannot be that
of R-Truth conditions for the sentences of a language. In his own paper
delivered at the Jerusalem conference, Dummett offers the following

observations of the tenability of meaning-theoretic realism:

> If meaning is use, that is, if the knowledge in which
> a speaker's understanding of a sentence consists must be
> capable of being fully manifested by his linguistic prac-
> tice, it appears that a model of meaning in terms of a
> knowledge of truth-conditions is possible only if we
> construe truth in such a way that the principle of
> bivalence fails; and this means, in effect, some notion
> of truth under which the truth of a sentence implies the
> possibility of <u>our</u> recognising its truth...It means that
> we cannòt operate, in general, with a picture of our
> language as bearing a sense that enables us to talk about
> a determinate, objective reality which renders what we
> say determinately true or false independently of whether
> we have the means to recognise its truth or falsity.   7

In the final sentence of this passage, we find Dummett tacitly extending his argument to include <u>semantic</u> realism, an extension we shall examine below. Of more immediate concern, however, are the implications of Dummett's remarks for his conception of <u>ontological</u> realism. In "WTM-II", he claims that, to the extent that we assume that a test for a particular property

> reveals an antecedently existing and independent state
> of affairs, we are adopting a realistic attitude towards
> the property or quantity in question...We may in fact
> characterise realism concerning a given class of state-
> ments as the assumption that each statement of that
> class is determinately either true or false.   8

This characterisation of <u>ontological</u> realism with respect to a given class of statements is further refined in the paper, "Realism", where Dummett attempts to survey the field of possible realist and anti-realist options. He begins by making it clear that he is concerned with debates over what we are terming ontological realism:

> It is clear that one can be a realist about one subject-
> matter, and not about another: though someone may have

a general inclination towards realistic views, it is plain
that there is no coherent philosophical position that
consists in being a realist tout court. This may be
expressed by saying that one may be a realist about cer-
tain entities - mental-states, possible worlds, mathemat-
ical objects - and not about others. But it seems pref-
erable to say that realism is a view about a certain
class of statements - for instance, statements in the
future tense, or ethical statements - since certain
kinds of realism, about the future or about ethics, do
not seem readily classifiable as doctrines about a realm
of entities. So, in every case, we may regard a realis-
tic view as consisting in a certain interpretation of
statements in some class, which I shall call the "given
class".   9

Realism, so construed, is, so Dummett maintains

a semantic thesis, a thesis about what, in general, ren-
ders a statement in the given class true when it is true.
The very minimum that realism can be held to involve is
that statements in the given class relate to some reality
that exists independently of our knowledge of it, in such
a way that that reality renders each statement in the class
determinately true or false, again independently of
whether we know, or are even able to discover, its truth-
value.   10

Realism for a particular class of statements thus requires that one

accept the principle of bivalence for members of that class. But Dummett

modifies his earlier account, in "WTM-II", by adding that acceptance of

bivalence is not sufficient for realism:

To have a realistic view, it is not enough to suppose
that statements of the given class are determined, by
the reality to which they relate, either as true or as
false; one has also to have a certain conception of the
manner in which they are so determined. This concep-
tion consists essentially in the classical two-valued
semantics: and this, in turn, embodies an appeal to the
notion of reference as an indispensable notion of the
semantic theory.   11

The foregoing characterisation of ontological realism as a "seman-

tic thesis" with the "minimal content" specified in the above passage suggests, however, that Dummett subscribes to the view that ontological realism for a given class of statements entails the ascription of R-Truth conditions to members of that class: in other words, it appears that, for Dummett, ontological realism for a class of statements entails what we are terming "semantic realism" for that class. But, as we have seen, Dummett also seems to think that his arguments against meaning-, theoretic realism extend to semantic realism (see the final sentence of the passage cited as footnote 7 above). This generates an apparent contradiction within Dummett's views:

A: The debate over meaning-theoretic realism is carried on at the level of the language as a whole, rather than in terms of particular classes of statements. So, if the arguments against meaning-theoretic realism extend to semantic realism, the success of the former arguments would seem to entail the rejection of semantic realism for the language as a whole. But, as just noted, Dummett seems to be committed to the following entailment: ontological realism for class C entails semantic realism for class C. And, in that case, the soundness of the arguments against meaning-theoretic realism would warrant a generalised ontological anti-realism. Furthermore, if the foregoing entailments hold, it would also be the case that, if we could establish ontological realism for any individual class of statements, we would thereby invalidate the arguments for meaning-theoretic anti-realism. And this is precisely the sort of possibility that Dummett seems to be entertaining in the following remarks on the subject of realism about the past:

> `I suspect, indeed, that if we are to discover the flaw
> in the general anti-realist contention that we can have
> no genuine notion of truth for our statements that
> transcends the kind of ground on the basis of which
> we are able to assert them, or the limit on the domain
> of application of this principle, we are most likely
> to discover it by studying this case.   12

B: The possibility, entertained in A, of an argument establishing a

generalised ontological anti-realism seems to run counter to Dummett's

already cited contention (see footnote 9 above) that "there is no

coherent philosophical position that consists· in being a realist tout

court", and that the debate between ontological realists and their

opponents is to be resolved on a case•by-case basis, Of course, it might

be pointed out that, in the passage cited, Dummett says nothing about

the coherence of being an anti-realist tout court, which is what is

canvassed in A. But, apart from the inherent implausibility of a gener-

alised ontological anti-realism, we have other textual grounds for

rejecting this proposal. In commenting upon his own earlier attempts

to provide a characterisation of ontological realism that would apply

to the various debates between realists and anti-realists, Dummett

asserts that

> there was...never any presumption that there would be any
> sound argument establishing, for all the cases simultan-
> eously, the correctness of a realist or of an anti-realist
> view...There could, indeed, be said to be a single higher-
> order issue: under what circumstances are we entitled to
> assume the principle of bivalence for some class of state-
> ments? The only answer that would yield the correctness
> of a realist view in every case would be the wildly im-
> plausible answer, "For every class of statements what-
> ever." On the other hand, the answer, "Only when we have
> an effective means of deciding the truth or falsity of
> each statement in the class," would yield the correct-
> ness of the anti-realist view in every disputed case.
> I have never thought it likely that so simple an ans-
> wer would prove to be tenable.   13

These remarks are interesting not only because Dummett evinces equal skepticism concerning the possibility of a general argument for ontological realism or for ontological anti-realism, but also because, in dismissing the latter possibility, there is no reference to the arguments against meaning-theoretic realism. Yet Dummett does not repudiate those arguments, and we have already noted that those arguments operate at the level of the entire language. This suggests that the entailments to which we appealed in A, as providing logical connections between meaning-theoretic realism and ontological realism, may in some way misconstrue Dummett's position. I shall attempt to substantiate this suggestion, and thereby remove the apparent inconsistency set forth in A and B above.

In the Preface to Truth and Other Enigmas, Dummett responds to Strawson's charge that the case for "the new anti-realism" rests upon an implausible attempt to generalise from mathematical intuitionism to discourse as a whole. Strawson claims that

> the new anti-realism starts from certain views on mathematical truth and mathematical discourse and seeks to generalise them to the extent of advancing counterpart views on discourse concerning the natural world...There is a certain initial air of paradox about this approach; for I suppose that at least part of the appeal of anti-realist views about pure mathematics lay precisely in the contrast between the content of that science and the subject matters of history, geography, natural science, and ordinary chat; that it lay precisely in the view that the notion of a realm of facts waiting to be explored, some parts of which might, indeed would, remain undisclosed or be irrecoverably lost sight of (a matter at best of speculation or uncertain inference) — that any such notion was quite improperly imported into mathematics from its natural home, viz. the natural world. 14

Dummett's response is **not** that Strawson is confusing what we have termed meaning-theoretic realism and ontological realism - the response that one might make if one upheld the logical independence of the two doctrines - but I think his response implies that Strawson is guilty of such a confusion. What Dummett challenges is the idea that disputes about ontological realism for a given class of statements are to be settled by appeal to what we might call a "metaphysical picture" - the sort of picture to which Strawson appeals in his talk of "a realm of facts" that justifies, through its presence, realism about the natural world, and undermines, through its absence, realism in pure mathematics. Strawson's conception of how disputes about ontological realism are to be resolved is, Dummett admits, a natural one:

> It looks at first sight as though, when we make state-
> ments about, say, the physical universe, there is a
> quite determinate objective, external reality, existing
> independently of us, which we are speaking about, and we
> can therefore safely take the content of our statements
> to be given by the way they are rendered true or false
> by that external reality, independently of whether we
> know, or are able to discover, their truth-value.   15

He argues, however, that the adoption of a particular "metaphysical picture" does not resolve the issue of ontological realism for a given class. Espousing the characteristically Platonist "metaphysical picture" of the ontological status of mathematical entities, for example, does not commit one to ontological realism in respect of pure mathematics:

> Just saying that natural numbers, existing independently
> of our thinking of them, have properties that exceed
> our capacity to characterise is of no help in confer-
> ring a meaning on expressions of our language; the mean-
> ing of sentences of our language can be given only by
> reference to procedures for allotting truth-values to
> them that we are able to carry out.   (Stress added)   16.

Dummett's point is that endorsing this sort of "metaphysical pic-
ture" of the nature of mathematical reality is insufficient to guaran-
tee ontological realism for the given class, because it is insuffic-
ient to guarantee that all sentences belonging to that class, as under-
stood by speakers of the language, satisfy the requirement of bivalence.
For, even if the "metaphysical picture" permits the ascription of deter-
minate R-Truth conditions to all sentences in the given class, this
will not yield bivalence for the class unless it is also the case that
such R-Truth conditions are properly ascribable to the sentences as
part of their meaning in the language of the speakers. But it is pre-
cisely this latter point that is challenged by arguments against meaning-
theoretic realism - arguments to which Dummett is presumably alluding
in the stressed section of the quoted passage. Thus arguments against
meaning-theoretic realism (and thereby, on Dummett's view, against
semantic realism) eliminate one way in which a realist stance towards
a given class of statements might be justified. But this will not elim-
inate the possibility of justifying such a realist stance altogether,
in so far as there are other ways of grounding the assumption of bival-
ence for the given class. It is by offering a ground for this assumption,
rather than by conjuring up a "metaphysical picture", that one argues
for a particular form of ontological realism:

> To justify a realist interpretation of statements of any
> kind, it was necessary to demonstrate that we had con-
> ferred on those statements meanings such as to yield
> a notion of truth, as applied to them, with respect to
> which each statement could be taken to be determinately
> either true or false, that is, under which the prin-
> ciple of bivalence held good.   17

The general question to be answered is: "Under what circumstances are we entitled to assume the principle of bivalence for some class of statements?"[18] If we reject the option of meaning-theoretic realism - whereby bivalence is guaranteed by R-Truth and the determinate nature of THE WORLD - then we must look to the sorts of truth-conditions that we _confer_ on particular classes of statements in order to determine whether realism in respect of a given class is justified.

In resolving our earlier puzzlement, it is important to clarify the place of the "realist metaphysical picture" in Dummett's account. For, we may recall, it is the use of the language of this picture in the characterisation of ontological realism in the paper, "Realism", that led us to attribute to Dummett the view that ontological realism for a given class entails semantic realism for that class. He claims, interestingly, that adherence to a particular "metaphysical picture" for an area of discourse is more a _consequence_ of taking a realist or an anti-realist stance towards a given class of statements than a _ground_ for such a decision:

> If one believes, with the platonists, that we have con-
> ferred on our mathematical statements meanings such as to
> render them all determinately either true or false inde-
> pendently of our knowledge, then one will find it natural
> to adopt the picture of a mathematical reality existing,
> fully determinate, independently of us. If, on the other
> hand, one believes, with the intuitionists, that the con-
> tent of a mathematical statement resides entirely in our
> ability to recognise what constitutes a proof of it, so
> that, when we lack an effective means of arriving at a
> proof or disproof, we have no right to declare it either
> true or false, one will prefer a picture according to which
> mathematical reality is constructed by us, or, at least,
> comes into existence only as we become aware of it.   19

One finds a similar point made in "WTM-II", where Dummett inquires

as to our requirements in a truth predicate. An essential ingredient in our concept of truth, he maintains, is a distinction between "the truth of what someone says and the grounds which he has for thinking it true:"[20]

> There is thus built into the concept of truth from the outset a contrast between the semantic and the pragmatic aspects of an assertion; truth is an objective property of what the speaker says, determined independently of his knowledge or his grounds for or his motives in saying it. [21]

This is not, Dummett stresses, in any way an answer to the question how we might justify ascribing to speakers a grasp of R-Truth conditions for the sentences in their language. But, he further maintains, there is a principle - principle C - which is partly constitutive of our notion of truth and which underlies attempts to construe truth as R-truth. Principle C states that "if a statement is true, then there must be something in virtue of which it is true." Dummett observes, in a manner reminiscent of his account of the place of "metaphysical pictures", that Principle C is essentially _regulative_ in nature:

> It is not so much that we first determine what there is in the world, and then decide, on the basis of that, what is required to make each given statement true, as that, having first settled on the appropriate notion of truth for various types of statement, we conclude from that to the constitution of reality. [22]

The point, as I understand it, is that the "correspondence to reality" of members of a given class of statements is a _consequence_ of our practice of ascribing determinate truth-conditions to the members of that class - of our adopting a realist stance twoards the class - rather than a ground that might _justify_ that practice.

Finally, we may return to Dummett's puzzling observation that ontological realism concerning the past might undermine the case for meaning-theoretic anti-realism. The challenge presented to the meaning-theoretic anti-realist by realism about the past is _not_ that it is an acceptable form of ontological realism (if indeed it is), but, rather, that it seems to be a form of ontological realism for which we cannot provide an anti-realistically acceptable account. In other words, we seem to lack an account of how, apart from through a grasp of R-Truth conditions, we might confer upon the class of statements about the past meanings capable of sustaining bivalence for that class. It is only if an intuitively acceptable form of ontological realism satisfies this further condition that it poses a serious challenge to the cause of meaning-theoretic anti-realism, as Crispin Wright points out in the passage cited earlier.

VI-b   Dummett's Anti-Realism: The Arguments

The essential task of a "theory of meaning", according to Dummett, is "to display the conventional significance of an utterance of any sentence of the language, what difference, as it were, is made to the world by the fact that it has been uttered."[23] Dummett also holds, as we have seen, that a theory of meaning for a language L must elucidate what it is to <u>understand</u> L - since meaning just <u>is</u> what is understood by the competent speaker - where "understanding" manifests itself in the practical ability to <u>use</u> the language, "to engage in a certain complex of activities involving the utterance of sentences."[24] One way in which a theory of meaning might discharge this obligation to the theory of understanding - a way which seems to be favoured by Dummett himself, if we go by such papers as "WTM-II" - is by ascribing to the competent speaker a body of <u>knowledge</u> whose articulation mirrors the articulation of the language itself. I shall term any conception of the proper form of a theory of meaning that endorses the preceding assumptions a "Dummettian" conception of such a theory. The further requirement that a theory of meaning be "full-blooded", if also endorsed, would make it incumbent upon a "Dummettian" theory of meaning to explain in what the language user's possession of the relevant body of knowledge consists, and how the possession of such knowledge issues in the ability to use the language.

Dummett maintains that any theory of meaning will have two components, which we may label M1 and M2. M1 provides an account of "the meaning of the individual words of the language and the principles governing the formation of phrases and sentences, in such a way as to determine

the specific content of every sentence of the language."[25] M1 will employ certain "central notions"; for example, in a realist theory of meaning, the central notion will be that of R-Truth. The task of M2 is to specify the connection between meaning, as explicated in M1, and the use of the sentences in L. In the case of a "Dummettian" theory, M2 will contain, inter alia, an explanation of the substance of the attribution, to competent speakers of L, of a knowledge of the meanings of expressions and sentences of L.

A theory of meaning might depart from the "Dummettian" conception in at least two ways. Firstly, it might violate the requirement that a theory of meaning be a theory of understanding - in the sense that meaning is what competent speakers understand - by permitting M1 to contain elements a knowledge of which is not ascribed to speakers by the theory of meaning as a whole. In this case, there is no prima facie reason why one should not combine meaning-theoretic anti-realism with semantic realism, to the extent that the notion of R-Truth is admitted as just such an element in M1. Secondly, a theory of meaning might respect the requirement violated in the first kind of departure from the "Dummettian" conception, but reject the assumption that the abilities of the language user are properly modelled by crediting the user with a body of knowledge of the sort proposed within the "Dummettian" model. In this case, as with the "Dummettian" conception itself, realism or anti-realism at the meaning-theoretic level seems to entail the same stance at the semantic level.

I shall defer until the end of this chapter consideration of the attempt to defend semantic realism by upholding a theory of meaning that departs from the "Dummettian" conception in the first of the above ways.

To the extent that we restrict our attention to theories of meaning
that either conform to the "Dummettian" model or depart from it in the
second of the specified ways, we may treat arguments against meaning-
theoretic realism as arguments against its semantic counterpart. In the
remainder of this section, I shall outline three arguments of the former
sort that Dummett has offered in the interests of rebutting semantic
realism. The first two arguments, which, following the literature, I
shall term the "Acquisition" and the "Manifestation" arguments, assume
a "Dummettian" conception of the theory of meaning. Dummett recognises,
however, that the realist can attempt to evade such arguments by rejec-
ting the underlying assumption that "the institution of language has
to be explained in terms of the ability of individual speakers, and
that this ability embodies knowledge of some kind."[26] He maintains that
such a realist is still vulnerable to a third argument, which I shall
term the "Use" argument. Let me now sketch the three arguments in turn;

1) The Acquisition Argument

Dummett's Acquisition argument against meaning-theoretic realism turns upon the following general principle: If we are to ascribe to competent speakers of L, as an elucidation of their understanding of L as manifested in their linguistic performance, knowledge of a certain kind, then it must be possible to give an account of how such knowledge is acquired in the process of mastering the language. Dummett argues that this principle cannot be satisfied if the knowledge in question is taken to be knowledge of R-Truth conditions for the sentences of L. Perhaps the clearest formulation of this kind of argument is the following:

> The anti-realist...maintains that the process by which we came to grasp the sense of statements of the disputed class, and the use which is subsequently made of these statements, are such that we could not derive from it any notion of what it would be for such a statement to be true independently of the sort of thing we have learned to recognise as establishing the truth of such statements. What we learn to do is to accept the truth of certain statements of the reductive class, or, in the case that there is no reductive class, the occurrence of certain conditions which we have been trained to recognise, as conclusively justifying the assertion of a given statement of the disputed class, and the truth of certain other statements, or the occurrence of certain other conditions, as conclusively justifying its denial. In the very nature of the case, we could not possibly have come to understand what it would be for the statement to be true independently of that which we have learned to treat as establishing its truth: there simply was no means by which we could be shown this. 27

The problem, in other words, is to explain how, given that we learn L by learning wher to assert, and when to refrain from asserting, sentences of L in response to the obtaining of recognisable conditions, we are able to associate with such sentences R-Truth conditions, whose ob-

taining may, ex hypothesi, transcend our recognitional capacities.

Elsewhere, Dummett considers and rejects two strategies that the realist might adopt in an attempt to surmount such difficulties. The first strategy utilises the "equivalence principle", which states that "any sentence A is equivalent in content to the sentence 'It is true that A'."[28] It might be maintained that we acquire the notion of R-Truth, and come to associate R-Truth conditions with sentences of L, by means of the equivalence principle. That the notion of truth so acquired is indeed a notion of R-Truth might be argued on the grounds that the sentences substitutable for 'A' in the equivalence schema will include sentences that are undecidable, e.g. sentences that quantify over infinite domains. Dummett argues, however, that this sort of strategy is not compatible with the realist's contention that the notion of R-Truth is a basic notion in the theory of meaning: for to endorse the strategy

> is in effect to suppose that we may acquire a mastery of
> the greater part of our language in advance of having any
> apprehension of the concept of truth: if we want to main-
> tain that what we learn, as we learn the language, is,
> primarily, what it is for each of the sentences that we
> understand to be true, then we must be able, for any
> given sentence, to give an account of what it is to know
> this which does not depend upon a presumed prior under-
> standing of the sentence; otherwise our theory of meaning
> is circular and explains nothing.   29

In other words, if we are to be able to understand the particular in-
stantiations of the equivalence principle through which the notion of
R-Truth is supposed to be acquired, we must already understand a meta-
language in which these instantiations are utterances. But, if the real-
ist theory of meaning is correct, our understanding of the meta-language

requires.that we already possess the notion of R-Truth.

The second realist strategy is to argue that we acquire the notion of R-Truth by imaginatively extending our grasp of the relation between <u>decidable</u> sentences and the <u>recognisable</u> conditions to which they are accountable. Dummett uses the following account of how we might acquire a realist understanding of sentences that quantify over infinite domains to exemplify this strategy:

> We gain our understanding of quantification over finite, surveyable domains by learning the procedure of conducting a complete survey, establishing the truth-value of every instance of the quantified statement. The assumption that the understanding so gained may be extended without further explanation to quantification over infinite domains rests on the idea that it is only a practical difficulty which impedes our determining the truth-values of sentences involving such quantification in a similar way; and when challenged is defended by appeal to a hypothetical being who could survey infinite domains in the same manner as we survey finite ones...In this way, we try to convince ourselves that our understanding of what it is for undecidable sentences to be true consists in our grasp of what it would be to be able to use such sentences to give direct reports of observation. We cannot do this: but we know just what powers a superhuman observer would have to have in order to be able to do it - a hypothetical being for whom the sentences in question would <u>not</u> be undecidable. And we tacitly suppose that it is in our conception of the powers which such a superhuman observer would have to have, and how he would determine the truth-values of the sentences, that our understanding of their truth conditions consists. · 30

Dummett is no more sympathetic to this realist strategy than he was to the first:

> [The realist account] works only by imputing to us an apprehension of the way in which those sentences might be used by beings very unlike ourselves, and, in so doing, fails to answer the question how we come to be able to assign to our sentences a meaning which is dependent on a use to which we are unable to put them. This difficulty faces any explanation of the meanings of certain

expressions which consists·solely of saying that we under-
stand those expressions by analogy with or extrapolation
from other expressions whose meanings we have come to grasp
in some more direct way. There is no way of distinguish-
ing such an account from the thesis that we treat cer-
tain of our sentences as if their use resembled that of
other sentences in certain respects in which it in fact
does not; that is, that we systematically misunderstand
our own language.   31

This response is interesting in that it rests, implicitly, on the

assumption that the answer to the canvassed question, "How can we come

·to be able to assign to our sentences a meaning which is dependent on

a use to which we are unable to put them?", is that we can't achieve

such a feat. Only if we make this assumption will we be entitled to

characterise the ascription to our sentences of a meaning not so dep-

endent on use - if this is indeed what the realist is committed to -

as a case of "systematically misunderstanding our own language." In

fact, the assumption in question is one that we have already encoun-

tered (see footnote 16 above) in Dummett's claim that "the meaning of

sentences of our language can be given only by reference to proced-

ures for allotting truth-values to them that we are able to carry out."

If we wish to further explore the motivation behind, and the implicat-

ions of, such an assumption, we must turn to the Manifestation argument.

2) The Manifestation Argument

The Manifestation argument charges that a realist theory of meaning that conforms to the "Dummettian" conception cannot satisfy certain basic adequacy conditions for such a theory of meaning. These conditions are that (i) a theory of meaning should be a theory of understanding, and (ii) a theory of meaning should be "full-blooded" rather than "modest". Dummett claims that the latter condition - which in effect proscribes a tacit appeal to speakers' knowledge of the metalanguage in an elucidation of the "knowledge of meaning" ascribed to them by the theory of understanding - must be satisfied if the theory of meaning is to provide the desired philosophical clarification of the operations of language. The former condition has already been discussed. Taken together, the two conditions impose upon an adequate "Dummettian" realist theory of meaning the requirement that it "supply an explanation of what it is to ascribe to someone a knowledge of the condition which must obtain for a sentence to be true,"[32] that is, in our terminology, the R-Truth condition of the sentence. Dummett claims that a realist theory cannot satisfy this requirement because it follows from the very nature of the enterprise of constructing a theory of meaning that the "explanation" mandated by the requirement must show how every element in the meaning ascribed to a sentence is capable of being manifested in use:

> A meaning-theory, being a theoretical representation of a practical ability, must not only say what a speaker must know in order to know the language, but in what his having that knowledge consists, that is, what constitutes a manifestation of it...If meaning is use, that is, if the knowledge in which a speaker's understanding of a sentence consists must be capable of being fully manifested by his linguistic practice, it appears that a model

> of meaning in terms of a knowledge of truth-conditions is
> possible only if we construe truth in such a way that the
> principle of bivalence fails; and this means, in effect,
> some notion of truth under which the truth of a sentence
> implies the possibility, in principle, of <u>our</u> recognising
> its truth.   33

Note that the justification for what we may term the "manifestation"

requirement - the requirement that the knowledge that a theory of mean-

ing ascribes to competent speakers must be capable of being manifested

in use - is essentially <u>methodological</u>: if the body of knowledge as-

cribed to the language user is intended to be "a theoretical represen-

tation of a practical ability", there can be no warrant for including

in that body of knowledge elements of meaning that are not exhibited

in the practice one is endeavouring to represent.

Let me now sketch in some of the details of the argument that

Dummett constructs on the basis of the above assumptions. The realist

protagonist is one who maintains that a competent speaker's understand-

ing of a language L is properly represented as a knowledge of R-Truth

conditions for the sentences of L. The question is whether the realist

can explain what such knowledge consists in, thereby satisfying the

above-mentioned adequacy conditions on a theory of meaning. Dummett

grants that the realist can discharge this obligation to the theory of

understanding in the case of sentences that a speaker can come to under-

stand through a verbal explanation: in such cases, we may credit the

language user with <u>explicit</u> knowledge of a truth-condition, manifested

in the ability to <u>state</u> that condition. The realist's problems begin

when we recognise that one cannot discharge the <u>entire</u> obligation to

the theory of understanding in this manner:

An explanation of this form obviously presupposes that the
speaker already knows a fairly extensive fragment of the
language, by means of which he can state the conditions
for the truth of the given sentence, and in terms of which
he came to understand it. It follows that, however large
the range of sentences of the language his understanding
of which can be explained in this way, this form of ex-
planation will not suffice generally. Since, in virtue of
the equivalence principle, to state the condition for the
truth of a sentence is simply to give the content of that
sentence in other words, an explicit knowledge of 'the truth-
condition of a sentence can constitute a speaker's
grasp of its meaning only for sentences introduced by
means of purely verbal explanations in the course of his
progressive acquisition of the language: it would,
notoriously, be circular to maintain that a speaker's
understanding of his language consisted, in general, of
his ability to express every sentence in other words,
i.e. by means of a distinct equivalent sentence of the
same language. His understanding of the most primitive
part of the language, its lower levels, cannot be explained
in this way: if that understanding consists in a know-
ledge of the truth-conditions of sentences, such know-
ledge must be implicit knowledge, and hence the theory of
meaning must supply us with an account of how that know-
ledge is manifested.   34

An "implicit" knowledge of the R-Truth condition of a sentence S

can be manifested by a speaker's holding S to be true just in case the

condition obtains. But such an account is available only when "the

condition in question is one which [the speaker] can be credited with

recognising whenever it obtains."[35] For if we allow that the truth-

condition of S can obtain unrecognisably, it is not apparent how a

speaker's ability to hold S true when certain recognisable conditions

obtain could also serve to manifest an implicit knowledge of the former -

i.e. of what it is for S to be unrecognisably true. But, in the case of

any sentences which are, in practice or even in principle,
decidable, that is, for which a speaker has some effective
procedure which will, in a finite time, put him into a
position in which he can recognise whether or not the
condition for the truth of the sentence is satisfied...,

> we may say that the speaker's knowledge of the condition
> for it to be true consists in his mastery of the proced-
> ure for deciding it, that is, his ability, under suitable
> prompting, to carry out the procedure and display, at
> the end of it, his recognition that the condition does,
> or does not, obtain.   36

The problem for the realist is that there are many sentences in natural language that are not effectively decidable in this manner, and our understanding of which cannot, in general, be characterised as an explicit knowledge of truth-conditions: "The existence of such senten- ces cannot be due solely to the occurrence of expressions introduced by purely verbal explanations: a language all of whose sentences were de- ctdable would continue to have this property when enriched by express- ions so introduced."[37] Dummett lists a number of features of natural language that permit the formation of sentences that are not in prin- ciple effectively decidable:

> The use of quantification over an infinite or unsurvey-
> able domain (e.g. over all future times); the use of the
> subjunctive conditional, or of expressions explainable
> only by means of it; the possibility of referring to
> regions of space-time in principle inaccessible to us.   38

Even if it is sometimes possible for speakers to recognise that the truth condition for an undecidable sentence obtains, or does not obtain, this is insufficient to warrant an ascription of implicit knowledge of the R-Truth condition of that sentence, as we noted above. In fact, Dummett argues, we are forced to conclude that a natural language will contain at least some sentences with respect to which we can justify ascribing neither explicit nor implicit knowledge of their truth condit- ions to competent speakers of the language:

> Whenever the condition for the truth of a sentence is one
> that we have no way of bringing ourselves to recognise as
> obtianing whenever it obtains, it seems plain that there
> is no content to an ascription of an <u>implicit</u> knowledge of
> what that condition is, since there is no practical abil-
> ity by means of which such knowledge may be manifested.
> An ascription of the knowledge of such a condition can
> only be construed as <u>explicit</u> knowledge, consisting in
> a capacity to <u>state</u> the condition in some non-circular
> manner; and that, as we have seen, is of no use here.   39

Thus a theory of meaning that attempts to elucidate linguistic under-

standing in terms of the knowledge of R-Truth conditions cannot satisfy

the requirement that such knowledge be ascribed only when fully accoun-

table to what can be manifested in linguistic practice.

3) The Use Argument

If Dummett's conception of the nature and import of the first two anti-realist arguments is relatively clear, the same cannot be said for what I am terming the "Use" argument. The general thrust of the latter is that a realist theory of meaning cannot satisfy the requirement - imposed as part of M2 - that the meanings ascribed to expressions and sentences be appropriately connected with the practice of using the language. The clearest formulation of such an argument is to be found in a paper by Dag Prawitz, "Intuitionist Logic: A Philosophical Challenge", upon which Dummett was the commentator.[40] However, as we shall see, Dummett and Prawitz seem to disagree in their interpretations of the significance of the argument. Dummett sees the argument as a counter to "non-Dummettian" realist theories of meaning, whereas, I shall suggest, Prawitz doesn't view it in this manner.

Prawitz begins by distinguishing two ways in which we might interpret the constraints imposed on a theory of meaning by the Wittgensteinian dictum that "meaning is use". Firstly, following Dummett[41], we might read the dictum as requiring that we observe the sort of "manifestation" requirement identified above - that is, that we ascribe knowledge of meaning to a speaker only to the extent that such knowledge is, or can be, manifested in use. Prawitz notes that, if we impose this sort of requirement on the theory of meaning, we invite the response that "although knowledge of the meaning of an expression must imply a certain use of the expression, the knowledge may not manifest itself in any finite totality of individual behaviour."[42] If we allow this objection, we undermine the Manifestation argument which rests upon a proscription upon the ascription of knowledge of meaning transcending

<u>manifestable</u> knowledge of meaning.

A second, and weaker, interpretation of the Wittgensteinian dictum - one which tacitly grants the force of the objection to the first inter- pretation - is formulated by Prawitz as follows:

> The main function of the concept of meaning is to explain communication. In general, we explain people's ways of using and reacting to utterances by, among other things, assuming that they know or do not know the meaning of the utterances. In an adequate theory of meaning, we should thus be able to deduce a certain use of the language from the meaning of its expressions. In particular, from an assumption that a person knows the meaning of an ex- pression should follow that he uses the expression in certain ways. 43

If this constraint is to be applied as a condition of adequacy on theo- ries of meaning, then it is necessary to ascertain <u>which</u> features of use are to be deducible from the meanings of expressions. Prawitz appeals to two main features of the use of an expression identified by Dummett: "the rules governing in what situations the expression is appropriately uttered and those governing the appropriate responses or expectations that follow from uttering the expression; schematically, the conditions for uttering the expression and the consequences of uttering it."[44] He then argues that a realist theory of meaning is un- able to satisfy even this weaker construal of the "meaning is use" dictum:

> A person cannot be said to know the meaning of an express- ion without knowing both aspects of its use. But if mean- ing is identified with truth conditions it is difficult to see how knowledge of meaning is to imply knowledge of use; in fact, it is even unclear what is to be meant by knowing a truth condition, since explicit knowledge cannot be demanded. 45

A similar line of reasoning occurs in the writings of Crispin

Wright. Wright maintains that "understanding a statement is a practical discriminatory ability: it is the ability to distinguish between circumstances which do and circumstances which do not justify the assertoric use of the statement."[46] He further maintains that an account of the truth-conditions of a sentence will not always enable one to distinguish between these two classes of circumstances:

> If the truth-conditions of a statement are not recognisable as such when fulfilled, then just to understand what they are - leaving aside the question of what such understanding might now be held to consist in - is not yet to know in response to what in one's experience the statement might be rightly used. [47]

Notice that both Prawitz's argument, as I have presented it, and Wright's argument address a realist theory of meaning of the "Dummettian" type - one that encompasses a theory of understanding that ascribes knowledge of R-Truth conditions to speakers of a language. The difference between the Use argument, so construed, and the Manifestation argument is that, while both trade upon the gap between R-Truth conditions and conditions that can always be recognised to obtain, they reason in different directions. The Manifestation argument proscribes an inference from the ability to regulate our use of a sentence in response to certain recognisable conditions, to a knowledge of R-Truth conditions that may obtain unrecognisably; the Use argument maintains that such a knowledge of R-Truth conditions, were it attainable, would not, by itself, enable us to correctly regulate our use of a sentence in response to certain recognisable conditions - for example, it may be incorrect to assert S in circumstances where S is in fact true, but where there are insufficient grounds for such an assertion.

In his commentary on Prawitz, however, Dummett offers a different reading of what the former's presentation of the Use argument is supposed to establish. He begins by suggesting that Prawitz questions the Manifestation argument because it rests upon, the "highly problematic" assumption that "the institution of language has to be explained in terms of the ability of the individual speakers, and that this ability embodies knowledge of some kind."[48] To reject this assumption is to deviate from a "Dummettian" conception of a theory of meaning in at least the second of the two respects distinguished above. But, as I suggested above, Prawitz's expressed reservation concerning the Manifestation argument seems to be much less radical than this His point is that, while continuing to endorse the "Dummettian" conception of the theory of meaning, one might reject the "manifestation" requirement, at least in its extreme form.

Dummett proceeds to outline Prawitz's argument in a manner broadly similar to my own presentation above, but he offers a different diagnosis of the force of the argument against realist theories of meaning:

> The advocate of a truth-conditional meaning-theory is in a dilemma. If he claims that a speaker actually knows the condition for the truth of each sentence, then he is open to [the Manifestation argument]. But if he denies that the speaker actually knows the truth-conditions of the sentences, it ceases to be in any way plausible that he [i.e. the realist] has a means of deriving the actual use of those sentences from the specification of their truth-conditions. [49]

Dummett's presentation of the first horn of the realist's dilemma suggests that the Use argument has no independent force against a realist who cleaves to a "Dummettian" conception of the theory of meaning, and I think that this misrepresents Prawitz's own claims for the argument.

The second horn of the dilemma, which accords with Dummett's reading
of Prawitz's designs, is more interesting. Dummett claims to find this
argument in section 6 of Prawitz's paper, though he qualifies his attrib-
ution and refers to the argument in its original form as "very com-
pressed". I must confess that, if the argument is indeed in section 6,
then it is so compressed as to entirely escape my apprehension. But the
argument is interesting in its own right, for it challenges the realist
who tries to uphold semantic realism independently of meaning-theoretic
realism to establish some connection between the property of R-Truth -
a property which some of our sentences supposedly determinately possess
and others determinately lack - and features of our actual linguistic
practice.

VI-c Dummett and His Critics

1) Devitt on Dummett

Critical response to Dummett's critique of semantic realism has generally focussed on only one of the three arguments distinguished in the previous section, namely, the Manifestation argument. It is not difficult to understand why this should be so. A rebuttal of the Manifestation argument would not only provide compelling grounds for _accepting_ semantic realism, but would also underwrite a temporising response to the other two arguments; whereas even a rebuttal of _both_ of the latter arguments would not, by itself, deliver either the first of these benefits or the appropriate analogue of the second. The Acquisition argument and the Use argument do indeed identify problems that any fully-developed realist theory of meaning must resolve, but to demonstrate that the semantic realist _can_ resolve these problems is no reason to espouse semantic realism given that, as seems obvious, these requirements on a theory of meaning can be satisfied by the semantic anti-realist also. To answer the Manifestation argument, on the other hand, is to show that our linguistic practice _does_ manifest a grasp of R-Truth - unless one's answer consists in the rejection of the "manifestation" requirement itself - and is thus, _ipso facto_, an argument for semantic realism. Furthermore, if it can be shown that we _do_ possess the concept of R-Truth and that we _do_ associate R-Truth conditions with the sentences of our language, it will follow that there must _be_ a satisfactory answer to the Acquisition argument, even if that answer eludes us, for the present (or even if it always eludes us, for that matter). As for the Use argument, it would seem that, even _without_ an answer to the Manifestation argument, the semantic realist can argue

that, while knowledge of the R-Truth condition for a sentence S is nec-essary for an understanding of S, it is not sufficient: what is also required is a knowledge of the conditions governing the assertion of S in circumstances where the R-Truth is not recognised to obtain.

In evaluating the critical response to Dummett, I shall, according-ly, focus upon proposed counters to the Manifestation argument, referr-ing to independent criticisms of the other arguments offered by writers under discussion when appropriate. There are, as we shall see, striking similarities between the sorts of strategies that realists have adop-ted in the face of the Manifestation argument, on the one hand, and the sorts of strategies - surveyed in the previous chapter - that have been employed to counter the "model-theoretic" argument. This should not sur-prise us, given that both arguments assail the "radically non-epistemic" nature of R-Truth, its capacity to transcend all possible verification. It will be helpful to begin by clarifying certain details of the Mani-festation argument, and a suitable foil for this endeavour is provided by Michael Devitt's paper, "Dummett's Anti-Realism".[50] After sketching Devitt's reading and criticism of Dummett's argument, I shall identify what I take to be the salient misinterpretations in Devitt's account.

Devitt maintains that Dummett's argument for semantic anti-realism - which, as we shall see, is, on Devitt's reading, a form of the Manifes-tation argument - "has a general form that is claimed to cover various 'realisms'...realism about common-sense physical entities, about scien-tific entities, and about the past [for example]."[51] Devitt further states that his aim is to defend one of these "realisms", namely, real-ism about common-sense physical entities, which, for the purposes in hand, he labels 'Realism' tout court. He offers the following character-

isation of Dummett's anti-realist argument:

> There are three premises in Dummett's argument for anti-realism:
>
> A, The Realism dispute is the dispute about whether statements have realist (evidence-transcendent) or only verificationist truth conditions.
> The statements in question here are, of course, statements containing words like 'stone', 'tree', and 'cat': "physical statements". Call the doctrine that these statements have realist truth conditions <u>Realist Truth</u>. For Dummett Realism <u>is</u> Realist Truth.
> B. The dispute about truth conditions is the dispute about whether the competent speaker's understanding is realist (evidence-transcendent) or only verificationist.
> It follows from B that if the understanding is only verificationist then Realist Truth is false. With A this leads to anti-Realism.
> C. The competent speaker's understanding is only verificationist.   52

Devitt finds reason to dispute each of the premisses of this argument. He maintains, firstly, that A "completely misconceives the Realism issue." Realism, as traditionally understood, might be characterised as the doctrine that "common-sense physical entities objectively exist independently of the mental,"[53] and it is apparent, he further maintains, that Dummett's argument is intended to undermine such a doctrine. But, Devitt argues, Realism so construed neither entails nor is entailed by Realist Truth:

> Realism says nothing about truth nor even about the bearers of truth, sentences and beliefs (except, perhaps, in its use of 'objective', the negative point that beliefs do not determine-existence) <u>Realism says nothing semantic at all</u>...Does Realist Truth entail Realism? It does not. Realism (as I have defined it) requires the objective independent existence of common-sense physical entities. Realist Truth concerns physical <u>statements</u> and has no such requirement: <u>it says nothing about the nature of the reality that makes those statements true or false</u>, except that it is objective.   54

Secondly, Devitt claims that B and C admit of two different inter-
pretations, depending upon whether or not one endorses what he terms
the "propositional assumption", the assumption that linguistic compet-
ence "consists (at least partly) in knowledge of truth (falsity) condit-
ions,"[55] where such knowledge is taken to be propositional in nature.
Interpreted in terms of the propositional assumption, B and C yield the
following:

> B1. The dispute about truth conditions is the dispute
>   about whether the competent speaker knows realist
>   (evidence-transcendent) or only verificationist truth
>   conditions...
> C1. The competent speaker knows only verificationist
>   truth conditions.   56

If, on the other hand, we reject the propositional assumption and con-
strue linguistic competence "as simply a practical ability," we get
the following:

> B2. The dispute about truth conditions is the dispute
>   about whether the sentences understood by the com-
>   petent speaker have realist (evidence-transcendent)
>   or only verificationist truth conditions.
> C2. The sentences understood by the competent speaker
>   have only verificationist truth conditions.   57

Given Devitt's additional claim that it is probably implicit in the
propositional assumption that "whatever truth conditions a sentence has,
the speaker will know it to have,"[58] B1 upholds, and B2 rejects, the
idea that a theory of meaning should conform to the "Dummettian" model.

Devitt claims that the B1 interpretation of B has the virtue of
explaining at least one puzzling matter, namely, "how...a semantic dis-
pute about the truth conditions of sentences [could] be a psychological
dispute about the competent speaker's understanding,"[59] for whatever

truth conditions a speaker <u>knows</u> a sentence to have must be truth conditions that it <u>has</u>. He maintains that the attribution of the propositional assumption to Dummett is "well based but not certain": the uncertainty stems from Dummett's use of "weasel words", e.g. his claim that linguistic ability can be <u>represented by</u>, rather than <u>identified with</u>, propositional knowledge. But, Devitt claims, the only arguments one can find in Dummett's writings for the propositional assumption are no more than "a travesty"[60], and there are good independent grounds for rejecting the assumption. One should not, therefore, take an argument incorporating B1 and C1 as a good reason for giving up Realist Truth:

> If the propositional assumption is false then a person committed to Realistic Truth should hold only to the view that the speaker understands sentences that <u>have</u> realist truth conditions. So it is quite beside the point to argue against him that speakers do not <u>know</u> realist truth conditions. Yet that is the central thrust of Dummett's verificationist argument against Davidson (...Davidson is open to the argument, of course, because he accepts the propositional assumption). Establishing merely that speakers do not know realist truth conditions casts no doubt on Realist Truth.   61

Thirdly, Devitt sketches a version of the Manifestation argument as an argument for C2. The gist of the argument is that the practical ability that constitutes a competent speaker's understanding of a sentence S is the ability to manifest the behaviour necessary to recognise whether or not the verificationist truth condition for S obtains. The argument then concludes as follows:

> (6) The speaker's understanding of S associates S recognitionally with verificationist truth conditions...
> (7) S has no truth conditions other than those associated with it recognitionally by the speaker's understanding.   Therefore
> (8) S has only verificationist truth conditions.   62

Devitt raises two kinds of objections to this argument. Firstly, he charges that it trades in objectionable forms of anti-holism. We can endorse the contention that "to have an ability is to manifest a part-icular sort of behaviour in the appropriate circumstances" only if we also endorse a behaviouristic account of uderstanding, one that is opposed by recent arguments for holism in psychology. And the notion of verificationist truth conditions "requires that there be a particular recognizable condition in which the belief expressed by S is conclus-ively justified," and this requires a commitment to an anti-holist[63] epistemology. Secondly, Devitt argues that we should not accept (7) above. To reject (7) is to reject the idea that every element in the meaning or "assertoric content" of an expression or sentence must be accountable, within the framework of a theory of meaning, to the theory of understanding - that is, it is to deviate, in the first of the ways specified above, from a "Dummettian" conception of the theory of mean-ing.

I shall postpone discussion of this final matter until later in this chapter. For the present, I shall examine Devitt's other criticisms of Dummett. It should be apparent, from our investigations in earlier sections of this chapter, that the criticisms of A and of B1 rest upon serious misreadings of Dummettian anti-realism and of the contribution of the Manifestation argument thereto. Devitt's "Realism" is what we have termed ontological realism concerning physical objects, while his "Realist Truth" is what we have termed semantic realism. Devitt's ob-jection to A, translated into our terminology, is that semantic realism and various forms of ontological realism are logically independent of one another. But, I have argued, Dummett makes no claim to the contrary.[64]

Devitt is far from being alone in conflating questions of semantic realism and questions of ontological realism as they arise in Dummett's writings, although Devitt's conflation is perhaps the most blatant. A similar confusion is, I think, operative in McGinn's contention that Dummett tacitly operates with two different notions of "realism".[65]

Similar considerations apply to Devitt's charge that Dummett's argument, if interpreted in terms of B1 and C1, misses the point against the proponent of Realist Truth who wisely abjures the propositional assumption. For, in the first place, I have already cited Dummett's own acknowledgement of precisely this limitation on the scope of the Manifestation argument - the argument at issue in Devitt's paper, as is clear from his formulation of the reasoning that supports C. Dummett's acknowledgement, it might be noted, occurs in a paper published two years prior to Devitt's article. Furthermore, to demonstrate that a realist theory of meaning that endorses the propositional assumption is untenable is not, I think, a trivial or a merely academic exercise. For not only is it the case that, as Devitt points out, Davidson's very influential work towards a realist theory of meaning takes the propositional assumption as a given; but also, as Devitt further acknowledges, "the propositional assumption about competence is a received wisdom of contemporary semantics."[66] Whether or not one can preserve Realist Truth by giving up the propositional assumption, the point of the Manifestation argument is to rebut less radical realist strategies, and nothing in those of Dummett's criticisms examined thus far gives us reason to think that the argument fails to realise this intention.

Given the above, it might seem that Devitt's further claim, that the C2 incarnation of the Manifestation argument fails to undermine a

semantic realism that eschews the propositional assumption, need not detain us. But we should ask whether any of his criticisms would also apply, _mutatis mutandis_, to the Cl incarnation of the argument. When we explore this matter, we uncover further misconceptions of the issues on Devitt's behalf. Note, firstly, that the charges of psychological behaviourism and epistemological anti-holism, if justified against the C2 version of the Manifestation argument, should also count against the Cl version: for the charges are grounded in premisses common to both versions of the argument - the "manifestation" requirement, that everything attributed to competent speakers by a theory of meaning, have a distinct behavioural display, on the one hand, and the assumption that we can associate conclusive verification or falsification conditions with individual sentences, on the other. Consider, then, Devitt's defence of the charge that the Manifestation argument traffics in an unacceptable psychological behaviourism. He maintains that

> recent work in the philosophy of mind decisively favors functionalism over this kind of behaviourism. The difference can be put like this, Behaviourism sees each mental state as a simple input-output function: to be in a mental state is simply to be apt to yield certain behaviour as output given certain stimuli as input. According to functionalism a mental state is not a simple input-output function: it is related to input and output by causal relations to other mental states, usually complicated relations. Competence with a word or sentence can no more be tied to a particular manifestation than can pain, love, belief, or bravery. 67

If these observations on the relative merits of behaviourism and functionalism are to bear upon the Manifestation argument, then the theory of understanding, as a component of the theory of meaning, must be construed as a _psychological_ hypothesis concerning the _mental states_

that should be ascribed to competent speakers of the language. That
Devitt views the theory of understanding in this way is further appar-
ent in his already cited puzzlement, concerning B, as to "how...a seman-
tic dispute about the truth conditions of sentences [could] be a psycho-
logical dispute about the competent speaker's understanding." But Dummett
explicitly rejects such a view:

> A theory of meaning [of the "Dummettian" kind] is not
> intended as a psychological hypothesis. Its function is
> solely to present an analysis of the complex skill which
> constitutes mastery of a language, to display, in terms
> of what he may be said to know, just what it is that
> someone who possesses that mastery is able to do; it is
> not concerned to describe any inner psychological mech-
> anisms which may account for his having those abilities.
> If a Martian could learn to speak a human language or a
> robot be devised to behave in just the ways that are
> essential to a language-speaker, an implicit knowledge
> of the correct theory of meaning for the language could
> be attributed to the Martian or the robot with as much
> right as to a human speaker, even though their internal
> mechanisms were entirely different.   68

In a footnote to his paper, Devitt refers to these remarks as
"mysterious", and charges that Dummett

> overlooks the distinction between psychological and
> physical mechanisms. A Martian or a robot that is phys-
> ically different from us might be psychologically the
> same. The remark...that a theory of meaning, which is
> for him a theory of understanding, is not a "psycho-
> logical mechanism"...is mysterious because if under-
> standing is anything it is (at least partly) psycho-
> logical.   69

Dummett's use of the Martian example is unfortunate in that it invites
precisely this sort of response. What he needs to maintain is that we
might be justified in attributing an implicit knowledge of the theory
of meaning to the Martian independently of a decision as to whether

a developed psychology would ascribe to the Martian the same mental states that it ascribed to us as an explanation of the linguistic abilities in question. I think that Dummett would indeed maintain this, for the same reasons that would lead him to reject Devitt's claim that understanding, as that notion is employed in the context of a theory of meaning, is psychological in nature. Let me expand on this.

We need to recall what a "Dummettian" theory of meaning is designed to do. The goal is to illuminate philosophical problems concerning language, and, more specifically, concerning linguistic meaning. The strategy is to focus on knowledge of meaning, which is what is meant by "understanding". But "knowledge of meaning" is itself practical knowledge, the ability to speak the language. The task of a theory of meaning is to provide a theoretical representation of this practical ability, a representation in terms of propositional knowledge:

> Mastery of a procedure, of a conventional practice, can
> always be so represented, and, whenever the practice is
> complex, such a representation often provides the only
> convenient mode of analysis of it. Thus what we seek is
> a theoretical representation of a practical ability. Such
> a theoretical representation of the mastery of an entire
> language is what is called by Davidson, and will be
> called here, a "theory of meaning" for the language;
> Davidson was, perhaps, the first to propose explicitly
> that the philosophical problems concerning meaning ought
> to be investigated by enquiring after the form which
> such a theory of meaning for a language should take.   70

A theory of understanding, within the confines of such a theory of meaning, is a theoretical analysis of a practical ability that is intended to subserve the purposes of a philosophical account of meaning. The "manifestation" requirement imposed upon the theory of understanding is, as mentioned earlier, of methodological origin. The rationale

for such a requirement is well expressed by Crispin Wright: one who
rejects or is unable to satisfy this requirement

> opens up a gap between what understanding really is
> and the practical abilities which he wants to see as
> issuing from it. Closing the gap and answering the man-
> ifestation challenge are one and the same thing. While
> it remains, there has to be scope for the question:
> What could possibly count as evidence for supposing
> that someone has the relevant conception when all poss-
> ible actual and hypothetical data, even if uniformly
> favourable to the hypothesis on the realist view, are
> consistent with its falsity?   71

To attribute to a speaker, as part of a theory of understanding, know-
ledge whose possession could never be manifested in the practical lin-
guistic abilities that one is attempting to analyse, can make no con-
tribution to a philosophical comprehension of the phenomena under inves-
tigation, Wright maintains. But if the "manifestation" requirement is
grounded in these sorts of methodological considerations, why should it
be vulnerable to the outcome of debates between proponents of behav-
iourism and functionalism in psychology? Indeed, why should we suppose
that the analysis of understanding, as dictated by the requirements of
a philosophical account of meaning, will in any interesting way map
onto the analysis of linguistic ability produced by a mature intention-
alist psychology?

It is clear how Devitt would answer these questions. He would
maintains that the very idea of an autonomous philosophical account
of the workings of language is chimerical. In the final section of his
paper, he argues that Dummett is mistaken in thinking that questions
in the theory of understanding are dialectically prior to a resolution
of the debate over Realism: our task, he claims, is

> to put together the most plausible comprehensive theory
> of the phenomena that confront us. Theories of language
> and understanding are only two among many·scientific
> theories that must be fitted into the comprehensive
> picture. Realism is an overarching empirical (scientific)
> theory or principle. [72]

This is familiar territory, and it raises once again fundamental questions about the relationship between philosophy and other areas of inquiry, and about the manner in which disputes over the nature of reference, meaning, and truth are to be arbitrated. In the present context, however, we may perhaps finesse these deeper issues by remarking the following points. Firstly, even if we grant Devitt's contention that a theory of understanding is, at least in part, a psychological theory, and that, as such, it should assign mental states to language-users on functionalist rather than on behaviourist grounds, it still remains to be shown that there _are_ functionalist grounds for assigning to speakers, as at least part of an account of their linguistic understanding, a knowledge of R-Truth conditions.

Secondly, our acquaintance with Putnamian themes should make us skeptical of the claim that the virtues of semantic realism can be tested by appeal to "Realism as an overarching empirical hypothesis", or, more generally, of the claim that questions in the philosophy of language are to be resolved by the methods of the sciences. Devitt thinks that Realist Truth is to be defended by an inference to the best explanation of the following sort:

> I think that the most promising of such inferences starts
> from Realism, together with some observations about the
> properties and relations of the objects the Realist
> believes in, and argues for Realist Truth and a non-
> verificationist doctrine of understanding...The major

> problem for this inference to the best explanation is
> in showing that the Realist needs an explanatory notion of
> truth, at all. However I think we can hope to show this need
> in order to explain learning and teaching. Showing this
> would complete the case against verificationism.   73

We may recognise, here, the strategy of Putnam's "sophisticated realist". As we saw in chapter two above, however, adherence to "sophisticated realism" as an empirical hypothesis cannot, by itself, resolve the dispute between semantic realists and their anti-realist opponents because "sophisticated realism" itself admits of a realist (externalist) and a verificationist (internalist) interpretation. And again, as we saw in chapter five above, an appeal to the ways in which language is embedded in our non-linguistic practice, including our scientific practice, cannot establish that sentences in our language have R-Truth conditions: and it is precisely such an appeal that Devitt has in mind when he talks of "some observations about the properties and relations of the objects the Realist believes in."[74] Finally, we noted in chapter four above that at least part of the moral of the "brains in a vat" argument as formulated by Putnam is that there are some questions about the nature and limits of linguistic representation whose solution requires the methods of the philosopher, not those of the scientist.

Thirdly, Dummett's contention that a theory of meaning is not a psychological hypothesis should not be read as a denial of the psychological reality of whatever it is that the theory of understanding ascribes to the competent speaker in its analysis of her practical linguistic abilities. If it is read as such a denial, one will be led to ask both (a) what possible relevance the Acquisition argument has for the debate over semantic realism, and (b) how a theory of meaning built

upon such a theory of understanding can furnish any insight, "philo-
sophical" or otherwise, into the functioning of language. What Dummett
is concerned to deny is that the articulation of linguistic ability
provided by the theoretical representation of that ability in a theory
of meaning is accountable to the articulation of mental states provided
by a psychological theory of language use. An analogy may be of assis-
tance here. Philosophers of psychology such as Fodor have argued for a
view they term "token physicalism"[75]. The token physicalist holds that
psychological states are indeed physical states, but rejects the infer-
ence that theory-construction in psychology should be subject to the
constraint that natural kinds proposed by an adequate psychological
theory be identifiable with natural kinds in the physical sciences, and,
more specifically, in neurophysiology. The argument for token physical-
ism, very roughly, is that the interesting kinds for psychology - the
classes of things for which we can frame interesting generalisations -
may be individuated functionally rather than in terms of their physical
realisation; and there is no reason to believe that such functionally
individuated kinds will stand in a type-type relation to natural kinds,
or sets of the same, as individuated by the physical sciences. The core
of this argument is that variance in our cognitive interests may result
in different segmentations of a given domain in different fields of in-
quiry.

We might, I think, offer a similar analysis of the relation be-
tween philosophical and psychological accounts of language as that
relation might be viewed by a Dummettian. The latter can espouse what
we might term "token psychologism" - the view that the "knowledge" as-
cribed to competent speakers by the theory of understanding possesses

psychological reality, but that our cognitive interests in developing
a theory of meaning differ from the interests that motivate the con-
struction of psychological theories of linguistic ability; and that
this divergence in cognitive interests proscribes imposing upon the
theory of meaning the constraint that it respect, in its articulation
of understanding, the segmentation of linguistic ability contained in
a mature psychology of language. But token psychologism will impose
certain constraints - most notably, that a theory of understanding may
not ascribe to speakers "knowledge" that could not be psychologically
realised. If the ability to use a language is something that is <u>learned</u>
(even if such learning requires the assistance of certain innate capac-
ities), this constraint would prevent a theory of understanding from
ascribing to speakers, as "knowledge", something which was neither learn-
able nor plausibly construed as "innate". This, I take it, is the
rationale behind the Acquisition argument as an argument against seman-
tic realism.

Let me conclude this discussion of Devitt by commenting briefly
on his charge that the Manifestation argument requires a commitment
to an unacceptable anti-holist <u>epistemology</u>. He claims that Dummett's
argument

> requires that there <u>be</u> a particular recognisable condit-
> ion in which the belief expressed by S is conclusively
> justified. In my view the best recent work in epistem-
> ology and the philosophy of science shows this to be an
> impossible requirement. The relation between worldly
> conditions and a justified belief is much more complic-
> ated than is presupposed by the requirement...No belief
> is <u>conclusively</u> justified. Each belief is tied loosely
> to a range of conditions in which, relative to other
> beliefs, it is justified in varying degrees.   76

The Manifestation argument, he maintains, "requires an unreconstructed positivist epistemology." These charges invite at least the following responses. Firstly, to the extent that we accept the proposed holistic epistemology and reject the idea that there are recognisable circumstances that provide conclusive justification, we surely augment rather than diminish the very difficulty for a realist theory of meaning that Dummett seeks to establish by means of the Manifestation argument. For it becomes even more obscure what might constitute a speaker's implicit grasp of the R-Truth conditions for a sentence S, given that we can no longer appeal to a recognitional grasp of the truth conditions of decidable sentences to illuminate, by analogical extension, what it is to grasp the R-Truth conditions of undecidable sentences.

Secondly, as Alexander George points out in a response to Devitt[77], it isn't clear that Dummett's notion of "conclusive verification" does violate holist strictures on "unrevisability". For, expanding on George's suggestion, there seems to be no reason why the sorts of "verification conditions" required for Dummett's purposes should not incorporate "ceteris paribus" clauses in the following fashion: a speaker's understanding of a sentence S consists, in part, in the the knowledge that, ceteris paribus, circumstances C conclusively justify the assertion of S - where the ceteris paribus proviso is to be filled out by reference to the obtaining of normal background conditions, the unpolluted nature of the senses, the efficient functioning of the apparatus, etc. Of course, a speaker's understanding of the sentences employed to flesh out the ceteris paribus clause will be subject to the same sort of analysis, as holism requires. But there is no reason to think that this

renders incoherent an account of linguistic understanding in terms of "conclusive verification" so construed - unless, that is, there is a more general incoherence attending any attempt to give an account of linguistic understanding within the framework of an holistic epistemology. But the latter possibility would surely be a problem for the holist rather than for Dummett.

Thirdly, it is open to the anti-realist, though not, perhaps, to one who follows Dummett's canonical presentations of the anti-realist case, to repudiate the attempt to do the theory of understanding in terms of the notion of "conclusive verification", opting instead for some form of "confirmation". A confirmationist account of understanding may be found in the writings of Putnam[78], Paul Horwich[79], Crispin Wright, and, in muted vein, in some of Dummett's own more recent offerings[80]. The following passage from Crispin Wright may serve to exemplify this type of theory of understanding:

> Grasp of the sense of a declarative sentence is an
> operational grasp of the distinction between states of
> information which justify its assertoric use and states
> of information which do not. The achievement of veri-
> fication is one state of the former kind. So where veri-
> fication is possible, it is quite unexceptionable to think
> of knowledge of truth-conditions, of this non-transcen-
> dent sort, as playing an essential role in grasp of mean-
> ing. But such knowledge will not be the whole of under-
> standing in any case where other kinds of informational
> state are conceived as criterial justification for the
> assertion of the sentence - where mastery of its use
> requires the ability to recognise the relevance of
> such states.   [81]

It is not apparent how, if at all, the substitution of this kind of account of understanding for a strict verificationist account would deprive the Manifestation argument of its bite upon semantic realism.

2) Anti-Realism and the Undecidable

I shall now turn to a more general consideration of the critical response to Dummett's arguments against semantic realism. In order to impose some structure upon our examination of the secondary (and tertiary!) literature, I shall distinguish, in broad terms, a number of strategies adopted by Dummett's philosophical opponents - strategies that closely mirror realist responses to the "model-theoretic" argument. Firstly, echoing attempts to rebut the latter by the proffering of "counter-examples", there are varied attempts to rebut Dummett's conclusions by demonstrating that ti ere are sentences our understanding of which can be adequately explained only by attributing to us a know-, ledge of R-Truth conditions. Secondly, echoing those who challenge Stage I of the "model-theoretic" argument, we have attempts to under-, mine the conclusion of the Manifestation argument - the rejection of meaning-theoretic realism - by disputing either the premisses or the mechanics of the argument itself. Finally, echoing those who challenge Stage II of the "model-theoretic" argument by appealing to "additional constraints" on reference, we have critics who maintain that meaning transcends understanding, and that semantic realism can survive the rejection of meaning-theoretic realism. I shall examine representative examples of the first of these strategies in the present section, and examine the remaining strategies in the two sections following.

It will be helpful to begin our discussion of the first strategy by considering an example offered by Colin McGinn in his paper, "Truth and Use".[82] McGinn starts by outlining the Manifestation argument as he understands it. The general requirement on a theory of meaning is that

"we must...be able, as theorists, to relate the ascribed semantic know-
ledge to some specific practical capacity, or capacities, to use the
language."[83] Given that at least some of this knowledge must be implicit,
Dummett's proposal. according to McGinn, is that we correlate knowledge
of truth conditions with

> a certain sort of recognitional capacity, viz. a cap-
> acity to recognise, or come to know, the truth-value
> of sentences. Understanding a sentence thus consists
> in an ability to determine its truth-value in some
> canonical way...[This capacity] consists in a dispo-
> sition, when prompted, to undertake a procedure which,
> in a finite time, terminates in a recognition, signalled
> by some overt gesture, that the sentence's truth con-
> dition is fulfilled or not fulfilled, as the case may
> be. It is because (actual) possession of such an effec-
> tive decision procedure for truth-value is a practical
> capacity that it provides the needed link between know-
> ledge of sense and use.   84

The problem for the realist, then, is that an implicit knowledge of
R-Truth conditions cannot be connected with use in this way in the case
of "undecidable" sentences: in the case of such sentences, "alleged
grasp of transcendent truth-conditions cannot be associated with poss-
ession of an effective method for deciding truth-value," and a know-
ledge of such truth conditions "does not confer upon speakers a prac-
tical grasp of a canonically direct route to establishing their truth-
value."[85]

McGinn proposes to evaluate Dummett's argument, so construed, by
examining its implications when applied to a hypothetical example. We
are asked to imagine a linguistic community, C, whose members are arbor-.
eally rooted on the north side of a mountain and are incapable of local
motion. In the language, L, spoken by members of C, there are proper
names for sheep, predicates that apply to sheep, quantifiers over the

domain of local sheep, and sentence operators, N and S, whose meaning is given thus: "On the north (south) side of the mountain, it is the case that...". McGinn claims that those sentences in L that are either quantified or of the form S(A) will be undecidable for members of C, because "given the actual capacities of members of C, it is perfectly possible that they should be systematically unable to determine the truth-value of these sentences, since they cannot transport themselves to the south side of the mountain."[86] At best, they can have indirect evidence, in the form of recognisable N-side states of affairs, for the truth-values of the sentences in question.

The conclusion that McGinn wishes to draw is that, if Dummett's anti-realist arguments are correct, the members of C should give up bivalence for L and espouse an anti-realist conception of the hidden side of the mountain. Here, as in his later remark that the purpose of his paper is to show that the Acquisition and Manifestation arguments cannot function as a generally valid argument-schema against various forms of realism, McGinn reveals the same sort of confusion of semantic and ontological realisms that we have already encountered in our discussion of Devitt. Semantic anti-realism entails ontological anti-realism for a given class of statements - given Dummett's characterisation of ontological realism - only if there is no (semantic) anti-realistically acceptable way in which we can confer bivalence upon the members of that class. It is therefore impossible to gauge the precise implications of semantic anti-realism for debates about the various forms of ontological realism until it is determined just which ways of conferring bivalence are acceptable from a semantic anti-realist perspective. But we may still bring out what I take to be the intended force of McGinn's

example if, as seems plausible, we grant that, in the case in hand, a
semantic anti-realist who argues in the manner described by McGinn clear-
ly is committed to an (ontological) anti-realist stance vis-a-vis the
class of "undecidable" sentences in question. The challenge which
McGinn's example seems to present to the semantic anti-realist is well
expressed by William Demopoulos in a critical discussion of McGinn's
paper:

> To manifest understanding of south-face sentences in lan-
> guage use is to be capable of recognising the truth of
> south-face sentences when their truth conditions obtain.
> But since the tree people can only recognise the truth
> conditions of north-face sentences, south-face sentences,
> if they are understood at all, must be reducible to the
> recognition of north-face truth conditions. Since they
> are not so reducible, south-face statements must not be
> understood, and this is clearly absurd.   87

What we want to say about McGinn's example, as Demopoulos implies,
is that members of C do understand south-face sentences, and that this
understanding does not consist in a grasp of north-face truth conditions,
but, rather, in a grasp of south-face truth conditions. But south-face
truth conditions are recognition-transcendent, on McGinn's account, and
we can therefore attribute to members of C the sort of understanding we
are intuitively inclined to attribute to them only if we reject the
Manifestation argument and recur to a realist theory of understanding.
Furthermore, McGinn maintains, if we reflect upon the failure of cer-
tain prima facie plausible realist counters to his Dummettian argument
for anti-realism concerning south-face states of affairs, it becomes
apparent that their failure results from conceding to Dummett that
"there can be no conception of a state of affairs that isn't a recog-
nitional conception."[88] But this, he argues, is simply an extension of

the "empiricist dogma" that our "ideas" cannot go beyond our impress-
ions. We should hold, rather, that we can acquire conceptions through
our attempts to _explain_ what is observed - for example, in the case
of the members of C, the ongoing disappearance and apparent reappear-
ance of individual sheep. Members of C, having _acquired_ in this way
conceptions of reality that transcend their recognitional capacities,

> can manifest their knowledge of transcendent truth-
> conditions...by interpreting the assertions of fellow-
> speakers as expressions of the very realist beliefs we
> have seen no good reason to deny them (where their assent
> to bivalence will be of obvious relevance to such an
> interpretation). It seems that this way of locating know-
> ledge of truth-conditions within the total activity
> of speech interpretation serves, unambitiously but sat-
> isfactorily, to relate conceptions of transcendent states
> of affairs to a practical linguistic capacity, to ac-
> tual use.   89

I shall postpone consideration of this suggestion as to how we
might manifest a knowledge of R-Truth conditions to the following sec-
tion, where I shall examine a more developed version of the same sort
of proposal by John McDowell. For the present, we may remark a more ob-
vious problem with McGinn's example. The problem, as Demopoulos has
also pointed out, is that McGinn's "anti-realist" analysis of the ex-
ample rests upon the ascription, to Dummett, of the thesis that a
theory of understanding is justified in crediting speakers with implic-
it knowledge of the truth conditions for a sentence S only if they
actually possess an effective decision procedure for S - that is, only
if S is _in practice_ decidable for such speakers. Underlying this inter-
pretation of Dummett, there lurks, as with Devitt, the assumption that
the motivation behind the "manifestation" requirement is ultimately
behaviourist in nature. We are to take the position of the meaning-

theorist as analogous to that of the radical translator, and we there-
fore require that any knowledge ascribable to language-users be strict-
ly accountable to behaviour in which such users can actually, and obser-
vably, engage, on this reading of the anti-realist's motivation. That
this reading, and McGinn's interpretation of the substance of the
"manifestation" requirement, are incorrect is clear, however, if we
consider not only Dummett's earlier-cited observation that the require-
ment is satisfied by sentences "which are, in practice or even in prin-
ciple; decidable" (stress added: see the passage cited as footnote 16
above), but also the following remarks:

> The understanding of any mathematical expression con-
> sists in a knowledge of the way in which it contributes
> to determining what is to count as a proof of any state-
> ment in which it occurs. In this way, a grasp of the
> meaning of a mathematical sentence or expression is
> guaranteed to be something which is fully displayed in
> a mastery of the use of mathematical language, for it
> is directly connected with that practice. It is not
> in the least required, on such a theory of meaning,
> that every intelligible statement be effectively decid-
> able. We understand a given statement when we know how
> to recognise a proof of it when one is presented to us...
> Our understanding of a statement consists in a capacity,
> not necessarily to find a proof, but only to recognise
> one when found.
> Such a theory of meaning generalises readily to the non-
> mathematical case. Proof is the sole means which exists
> in mathematics for establishing a statement as true:
> the required general notion is, therefore, that of veri-
> fication. On this account, an understanding of a state-
> ment consists in a capacity to recognise whatever is
> counted as verifying it, i.e. as conclusively estab-
> lishing it as true. It is not necessary that we should
> have any means of deciding the truth or falsity of the
> statement, only that we be capable of recognising when
> its truth has been established. The advantage of this
> conception is that the condition for a statement's being
> verified , unlike the condition for its truth under the
> assumption of bivalence, is one which we must be cred-
> ited with the capacity for effectively recognising when
> it obtains; hence there is no difficulty in stating what
> an implicit knowledge of such a condition consists in -

once again, it is directly displayed by our linguistic
practice.   90

It should be apparent from this passage that the rationale for the
"manifestation" requirement - the requirement, alluded to in the final
sentence, that ascriptions of implicit knowledge be accountable to prac-
tical display - cannot be "behaviourist" in the manner suggested above.
For, where we lack an effective decision procedure, there may be no dis-
play in actual practice of our capacity to recognise a verification of
a given statement, yet Dummett allows that in such cases this capacity
may still be attributed to us by a theory of understanding, and we may
properly be said to possess an implicit knowledge of conditions that
are in practice unrecognisable. Clearly, then, what the "manifestation"
requirement proscribes is the ascription to language-users of an im-
plicit knowledge of conditions that are in principle unrecognisable.
And the reason for this proscription is, I have already suggested,
methodological in nature - it can serve no possible explanatory purpose
to ascribe to speakers, as part of a philosophical account of the
kings of language, knowledge that could not manifest itself in any
possible practice.

Once we correct McGinn's misconstrual of the import of the "mani-
festation" requirement, there seems to be no difficulty giving an acc-
ount of the "mountain-dwellers" example that both conforms to our intu-
itions and is acceptable to the semantic anti-realist. To cite Demopou-
los once more, we can attribute a non-reductive understanding of south-
face sentences to McGinn's tree-people "if they are capable of formul-
ating non-circular canonical descriptions of south-face truth condit-
ions...[which] satisfy the further constraint that they be associated

with <u>finitary</u> empirical 'proof' procedures; whether or not the tree-people can carry out the procedures is irrelevant."[91] Only if we interpret "decidability" as <u>in practice</u> decidability does McGinn's example pose a challenge to the semantic anti-realist.

The preceding discussion draws upon a distinction between what is decidable (or recognisable) <u>in principle</u> and what is decidable (or recognisable) in <u>practice</u>. The intelligibility and import of such a distinction has not gone unquestioned in the literature.[92] Let me briefly cite one strand in the debate on this matter. In the context of presenting the Manifestation argument as an argument for semantic anti-realism, Crispin Wright maintains that an account of linguistic understanding should restrict itself to circumstances "which we can actually (or at least in principle) recognise to obtain", and he glosses this remark as follows"

> By "in principle" possibilities, I intend roughly: feats in practice possible for a being with the same range of abilities as we, but possessing them to a greater though still finite degree. Whether anything turns on this distinction, from the point of view of a prospective anti-realist, is an unsolved, and fundamental, question.   93

In a commentary on Wright's paper, Strawson suggests that until a more extensive accounting is given of these sorts of matters, the anti-realist challenge to semantic realism must be regarded as inconclusive:

> .We need to know <u>at least</u> what is to count as falling within the range of "recognisable situations", what is to count as "conclusive verification", whose capacity in fact or in principle to do the recognising is in question, what importance, if any, to attach to the disjunction "in fact or in principle", and what "in principle" means.   94

Wright's response to Strawson is similar, in spirit at least, to my response, in chapter five above, to Devitt's charge that analogous unclarities infect the notion of an "Ideal" theory. Wright acknowledges that the questions raised by Strawson are important, especially if one intends to develop a semantic theory in which the notion of verification is "cast...in the role traditionally played by the notion of truth." He argues, however, that the force of Dummett's arguments against semantic realism does not depend upon prior clarification of the obscurities noted by Strawson. For, however these matters are resolved, it will surely be the case that there are some sentences that are not "conclusively verifiable", some sentences "for which the claim, 'I have conclusively verified S', will be taken, ceteris paribus, to betray a misunderstanding of S."[95] Or, in terms of the distinction with which we began, however we draw the line between those statements that are, and those that are not, "in principle decidable", we can assume that the latter class will not be empty. But the meaning-theoretic realist is committed to giving an account of our understanding of all statements, including those that are "in principle undecidable", in terms of the knowledge of their R-Truth conditions. As long as there are some "in principle undecidable" statements, therefore, the realist must still confront the "manifestation challenge" to show how, given the methodological constraints governing the construction of a theory of meaning, we are to justify attributing to speakers an implicit knowledge of truth conditions that are in principle unrecognisable. Thus, if we follow Wright, we can maintain that the anti-realist challenge need not await a clarification of the distinction between "in practice" and "in principle" decidability.

In fact, it is not difficult to come up with statements that are plausible candidates for membership in the class of "in principle unde-cidable" statements however. (within reason) we choose to precise our definition of that class. Indeed, realists themselves are only too happy to volunteer examples, when it serves their philosphical purposes. Thus we find Strawson himself arguing[96] that the anti-realist cannot give an adequate account of our understanding of certain classes of statements -- whose truth conditions are "necessarily verification transcendent". His chosen examples are (i) statements about the sensations experienced by other persons, and (ii) statements about the relatively remote past - remote enough to outstrip the reach of anyone's memory. I shall confine my attention to Strawson's presentation of the first kind of example, and to the responses that Dummett and Wright offer in the face of it. Similar considerations apply, however, to the second type of example.

. Strawson asks us to consider a sentence such as, "John is in pain", as spoken of John by another person. He attributes to Wright the follow-ing account of what the realist wishes to say about such a case: such a sentence has a truth condition which is necessarily verification trans-cendent from the speaker's point of view, but a knowledge of which governs the speaker's use of the sentence in question and constitutes his grasp of the sentence's meaning. Strawson asks whether we can give an adequate account of our understanding of such sentences - granting, which seems indisputable, that we do understand them - without endor-sing the realist account. He argues that there are three possible anti-realist construals of the relationship between the meaning of such sen-tences and their R-Truth conditions, and that none of these options is satisfactory. If we use "J" to symbolise the sentence, "John is in pain",

we may characterise Strawson's three anti-realist options as follows:

(A) We can deny that "J" _has_ a truth condition: "It is just a right thing to say in a certain range of circumstances, and knowing its meaning is practical knowledge of what those circumstances are."[97]

(B) We can identify the truth conditions of "J" with the _evidence_ that warrants its assertion, that is, the "circumstances" alluded to in (A) - thus to know the meaning of "J" is to have a practical grasp of these truth conditions.

(C) We can retain the idea that "J" has R-truth conditions, but divorce such conditions from the _meaning_ of "J", save in the case of John him- self: "Anyone else can understand the sentence without knowing what its truth-condition is."[98]

Strawson rejects (C) as counter-intuitive - because it entails that one could have a full knowledge of what a sentence meant, yet be ignorant of what one was saying in uttering the sentence assertorically - and (B) as unacceptably reductionist. The problem with (A), which he regards as the least implausible option, is that, if "J" lacks a truth condition, we can give no account of what a speaker is _doing_ in utter- ing "J", i.e. in using the sentence assertorically, Strawson claims. Thus, he concludes, the only adequate account is the realist one:

> The correct and only plausible account of the whole matter is...[that] it _is_ part of what it is now fashion- able to call our general _theory_ of the world that we re- gard other people as subject to roughly the same range of sensations as we are painfully or joyously or indiff- erently aware of in ourselves; and it is in no way con- trary to reason to regard ourselves - as in any case, we cannot help doing - as justified in certain circum- stances in ascribing to John a particular state of feeling which we cannot in the nature of the case experience ourselves; and his being in which is there- fore, if such is the standard invoked, necessarily

verification-transcendent.

Both Dummett and Wright have responded to Strawson's argument, and their principal objections are essentially the same. Furthermore, these objections would seem to extend to any other examples the real- ist might offer purporting to demonstrate that the anti-realist cannot adequately account for our understanding of statements that are "in principle undecidable". Before sketching the main line of response to Strawson, let me briefly outline an additional argument of Dummett's that addresses specific features of Strawson's chosen example.

Dummett traces the appeal of a realist account of our understanding of sentences such as "J" to the following considerations. We find quite unacceptable a behaviourist interpretation of such third-person sensation-ascriptions - an interpretation that equates the sense of such ascriptions with the obtaining of certain behavioural evidence. For we want to say that "J" may be true even in the absence of pain- behaviour (John may be paralysed or acting stoically), and may be false even in the presence of paradigmatic pain-behaviour (he may be a good actor). We therefore conclude that the sense of the sentence - that which is grasped by one who understands it - must be given by truth conditions that go beyond such evidence, namely, by its being with John as it is with us when we are in pain. But since it either is or is not with John as it is with us when we are in pain, the sentence in ques- tion must be determinately true or false, though we may be unable to recognise its being true or being false. But, in this case, it seems that our understanding of third-person sensation-ascriptions consists in our grasping their R-Truth conditions - in which case, it appears

that we require a realist theory of understanding for at least one class of statements.

A more general problem with this argument, to which we shall turn shortly, is that it rests on the implicit assumption that we have to choose between realism and behaviourism. A more specific failing, however, as Dummett points out, is that the realist account for which we are supposed to opt requires that we disregard Wittgenstein's critique of private ostensive definition and assume that one's understanding of the meaning of the term 'pain' is ultimately derived from reflection on one's own private experience. The burden of Wittgenstein's argument, however, which Dummett endorses, is that, even when the term 'pain' is employed in first-person ascriptions, "the supposedly contingent connections with pain-stimuli and pain-behaviour are in fact essential to the employment of the word."[100]

Both Dummett and Wright take the principal failing of Strawson's argument to be its illegitimately restricting the options open to the anti-realist. Wright makes this point by identifying an implicit assumption that underlies Strawson's rejection of both option (A) and option (C), namely, the assumption that "what is stated by an assertoric use of 'John is in pain' has to be explained by reference to necessary and sufficient conditions for that sentence's truth," and that, as a consequence, "only sentences that have been assigned determinate truth-conditions can be used to state anything!"[101] In the case of option (C), this assumption is operative in ruling out the possibility that a speaker who possesses an anti-realist understanding of the sentence, "J", can thereby know what she is saying in asserting the sentence. In the case of option (A), it operates in the legitimation of the inference, from

the claim that "J" lacks R-Truth conditions, to the conclusion that
"J" cannot be used to make an assertion. Dummett makes the same sort
of point, albeit in a manner that might seem to run contrary to Wright's
observations. He claims that Strawson assumes that, to the extent that
our account of the meaning of sentences like "J" refers only to grounds
of assertion that fall short of being <u>conclusive</u>, we must deny that
such sentences are properly characterised as being true or false, and
thus deny that they <u>have truth conditions</u>: such an assumption is, he
maintains, implicit in Strawson's charge that the anti-realist must
hold that sentences like "J" lack truth conditions. In characterising
Strawson's "charge" in this way, Dummett is presumably referring, in a
shorthand fashion, to the general import of Strawson's discussion of
the three anti-realist options - namely, that, given the patent inade-
quacy of (B) and (C), the anti-realist has no alternative but to dis-
pense altogether with the notion of truth conditions in giving a seman-
tic account of "in principle undecidable" sentences. Dummett's point is
<u>not</u> that the anti-realist can, contrary to Strawson's assumption, incor-
porate R-Truth conditions into his theory of meaning - this would be a
self-defeating strategy - but, rather, that there is a notion of "truth-
conditions" available to the anti-realist in terms of which the asser-
toric use of the problematic sentences can be explained in an anti-
realistically acceptable way.

What Strawson overlooks in his inventory of anti-realist options,
so it is claimed, is a position that Wright labels option (D), and that
both he and Dummett attribute to the later Wittgenstein. Option (D), as
applied to the sentence "J", differs from (A) in that it assigns a
truth condition to the sentence, from (B) in being non-reductive, and

from (C) in holding that an understanding of the sentence does involve possessing a conception of what it is for the sentence's truth conditions to obtain. It differs from Strawson's "realist" account in that: (i) it maintains that the truth conditions of the kinds of sentences in question admit of only trivial formulations, of the form, "'J' is true iff J" - Dummett alludes, here, to Wittgenstein's espousal of a redundancy theory of truth, and (ii) the requirement that a theory of meaning explain what knowledge of the meaning of a sentence - in this case knowledge of its "truth-conditions" - involves is to be met by giving a reductive account of this knowledge, in terms of possessing the practical skills necessary for the appropriate use of the sentence. Dummett spells this out as follows:

> That involves knowing that pain behaviour, or the pre-
> sence of an ordinarily painful stimulus, is normally a
> sufficient ground for an ascription of pain, but one
> that can be rebutted, in the former case by the clues
> that betray the shammer or by subsequent disclaimer;
> learning the symptoms of inhibiting the natural mani-
> festations of pain, and the limits beyond which this
> is impossible; knowing the usual connection between
> pain and bodily conditions, and the sort of cases in
> which the connection may be broken; and so on. To know
> these and similar things is, on Wittgenstein's account,
> just to know what "John is in pain" means; and for one
> who knows this, there need be no more informative answer
> to the question what makes the sentence true than,
> "John's being in pain".   102

Wright puts the point this way:

> "Grasping the truth conditions" is thus constituted by
> possession of a complex discriminatory skill exercised
> in response to public circumstances. And it is a danger-
> ous error to think of it, as does the realist, as some-
> thing ulterior, the formation of a conception of some-
> thing inaccessible, a conception which informs and governs
> both exercise of the relevant skill and all other as-
> pects of the use of the sentence, e.g. in inferences.   103

Of course, the realist will resist the attempt to label the postulation that a competent speaker's "grasp of truth conditions" is something that "lies behind" and governs overt linguistic performance, a "dangerous error". Indeed, one of Strawson's charges against Wright is that he fails to provide any argument that undermines the legitimacy of such a postulation. The realist can readily grant that "grasp of the sense of a sentence cannot be displayed in response to unrecognisable conditions," while maintaining that this in no way establishes that speakers' responses to recognisable conditions

> can never be governed by a certain kind of conception -
> a conception of a state of affairs, of a condition of
> truth, which, for one reason or another, in fact or in
> principle, is not, or is no longer, or is not for the
> speaker, accessible to direct observation or memory.   104

Wright's response to this charge, as we have seen, is that the realist's postulations are methodologically unsound, in that they can serve no possible explanatory purpose in furtherance of the goals of a theory of meaning. But might it not be claimed that the desideratum that speakers should understand what they are saying when they engage in the assertoric use of "in principle undecidable" sentences provides such an explanatory role for the postulation of a knowledge of R-Truth conditions within a theory of understanding? Is Strawson's "illegitimate" assumption that an anti-realist cannot give an account of what is stated by such sentences really illegitimate? Wright suggests that the intuit- ive appeal of a negative answer to this question rests upon the follow- ing line of reasoning. We feel that "if all someone possesses is prac- tical knowledge of criteria for the assertion and denial of a sentence, and knows of no truth condition for it, then he cannot be said to know

what, if anything, it states," because we assume that the "practical
knowledge" in question is simply the knowledge that the obtaining of
certain conditions which fail to provide conclusive verification (fal-
sification) of a sentence S, provides the best possible grounds for
asserting (denying) S. Wright maintains, however, that this conception
of the practical knowledge attributed to competent speakers by an anti-
realist theory of understanding is too narrow. We should also credit
the speaker with an operational grasp of

> just how things have to develop in any state of infor-
> mation warranting the assertion, or denial, of the sen-
> tence, in order for that situation to cease to obtain.
> In fact, of course, provided someone has practical know-
> ledge of all the types of situation criterially warrant-
> ing the assertion, or denial, of an S of the relevant
> kind, he would have to have practical knowledge of all
> the "overturning conditions" also. It seems to me,
> therefore, that the initial plausibility of the argu-
> ment turned on implicitly interpreting the knowledge
> possessed by the trainee as not being of that extent.   105

Wright also addresses a more general misunderstanding of the import
of semantic anti-realism that underlies our reluctance to accept that
an anti-realist understanding of an "in principle undecidable" sentence
can amount to a knowledge of what that sentence says, what it can be
used to assert. Strawson claims that to embrace anti-realism requires
that we reject certain intuitively acceptable features of our "general
theory of the world"; and, extending Strawson's point, it might be ar-
gued that a commitment to meaning-theoretic realism permeates the very
ways in which we think about and describe the workings of our language
- not the least example of which is our ascription of determinate truth
or falsity to "in principle undecidable" sentences. We have just examin-
ed the suggestions of Dummett and Wright concerning how an anti-realist

semantics might accomodate the latter feature of our practice. Wright argues that similar strategies allow the anti-realist to preserve other aspects of Strawson's "general theory of the world". The anti-realist should not <u>challenge</u> those elements in our prephilosophical descriptions of our linguistic practice that seem to offer succour to the cause of meaning-theoretic realism; rather, he should attempt to show that, properly interpreted, the "realist" descriptions are merely expressions of those very features of our linguistic practice that the anti-realist's theory of understanding comprises. This strategy, as prosecuted by Wright, resembles Dummett's treatment of the "realist rhetoric" of ontological realism, as discussed in section VI-a above.

Wright focusses on two aspects of our "general theory of the world" cited by Strawson. Firstly, there is our practice of taking others to be susceptible to the same range of sensations as ourselves. This, Wright maintains, should be viewed as "an expression of the existence of a communal vocabulary of sensation, any element of which is applicable to any of us on the basis of communally acknowledged criteria."[106] Secondly, there is our practice of taking ourselves to be justified in certain circumstances in ascribing to others states of feeling that are "necessarily verification transcendent" as far as we are concerned. This, according to Wright, is an expression of "the essential <u>defeasibility</u> of other-ascriptions of sensations, the fact that any state of information which warrants such an ascription can always coherently be envisaged as being added to in such a way that the resulting state of information no longer does so."[107] As for our "privileged access" to our own sensations; this is not to be taken as a consequence of our associating R-Truth conditions with first-person sensation-statements,

but, rather, as an expression of the fact that the "grammar of sensat-
ion" doesn't permit that a person should both <u>understand</u> a particular
self-ascription of sensation and be <u>mistaken</u> in that ascription. Wright
draws the general conclusion that "here, as in other areas, we rather
too readily tend to assume that opposition to realism is bound to be
radically revisionary of our ordinary theories and linguistic prac-
tices."[108]

3) Manifestability and "the Epistemological Roots of Anti-Realism"

The second sor⁴ of strategy adopted by defenders of semantic realism is to argue, by reference to specific features of Dummett's arguments for anti-realism, that the latter fail to undermine meaning-theoretic realism, and that one can therefore preserve semantic realism by constructing a realist theory of meaning around a realist theory of understanding. I shall restrict my attention to the most common form of this strategy in the literature, namely, the claim that the Manifestation argument presupposes certain restrictions upon the ways in which linguistic understanding can be manifested, that these restrictions are unjustified, and that, once these restrictions are waived, it is apparent that the linguistic practice of competent speakers <u>does</u> manifest a realist understanding of the language. I shall begin by sketching a rudimentary version of this argument by Simon Blackburn. Then, after surveying Crispin Wright's response to Blackburn, I shall outline a more developed version of the argument by John McDowell. I shall conclude by examining McDowell's contention that the anti-realist is not entitled to the sort of strategy that Wright adopts against Blackburn, and, by implication, against McDowell himself.

Blackburn believes that Dummett's reasons for imposing the "manifestation" requirement on the theory of meaning are revealed in the following passage:

> The reason is that the meaning of a statement consists
> solely in its role as an instrument of communication
> between individuals, just as the powers of a chess-piece
> consist solely in its role in the game according to the
> rules. An individual cannot communicate what he cannot
> be observed to communicate...   109

The problem with the "manifestation" requirement, according to Blackburn, is that "manifestation is primarily a relational matter:"[110] in other words, what A can manifest to B depends, in part, upon the sorts of resources that B, as "manifestee", can draw upon. If we assume that the only resources available to B are her powers of passive observation which allow her to register the empirical circumstances in which A's utterances are produced - if, that is to say, we place B in the position of a Quinean "radical translator" - then it will clearly be impossible fo A to manifest to B an understanding of things which might be taken to lie outside the range of such powers - for example, the past, or the sensations of other persons. But why, Blackburn asks, should we adopt such an impoverished conception of the resources available to the manifestee? What is at issue, he claims, is not so much what can be manifested in the sense of being _displayed_ to the passive observer, but, rather, what B can _understand_ or _know_ about A's understanding of the language given A's linguistic performance. And, in this case, "the question of whether the meaning may be the function of something lying beyond observation will hinge upon what we say about our knowledge of such matters."[111] Thus, Blackburn concludes, the Manifestation argument constitutes a valid challenge to semantic realism only if supplemented by an appropriate theory of knowledge; and, he implies, the sort of epistemological theory that the anti-realist requires is not likely to be forthcoming.

Wright argues that Blackburn misses the point of the "manifestation" requirement. The anti-realist case does not rest on the assumption that the relation between manifester and manifestee can appropriately be modelled on that between a radical translator and a native speaker. Rather,

we can run the Manifestation argument even in the sorts of circum-
stances that should be most favourable to the realist cause, given
Blackburn's contentions - situations in which both A and B are native
speakers of the same language. In such a situation, how is A to mani-
fest to B that he possesses a recognition-transcendent conception of
the truth conditions of a sentence such as "J"? As the Manifestation
argument points out, the problem is that everything in the exemplary
use of this sentence by A is consistent with A's having no such con-
ception, "but only the complex of recognitional and inferential skills
which the anti-realist claims constitute my understanding of the sen-
tence."[112] Wright stresses that the problem is not the skeptical one
that "any finite segment of my behaviour may take a form consistent
with supposing that I understand the sentence when in fact I do not."
Rather, to quote again a passage cited earlier, the point is that

> the realist's account opens up a gap between what under-
> standing really is and the practical abilities which he
> wants to see as issuing from it. Closing the gap and
> answering the manifestation challenge are one and the
> same thing. While it remains there has to be scope for
> the question: what could possibly count as evidence for
> supposing that someone has the relevant conception when
> all possible actual and hypothetical data, even if uni-
> formly favourable to the hypothesis on the realist view,
> are consistent with its falsity.   113

The objection to the realist analysis, as we have already stressed,
is the methodological one that no possible explanatory purpose can be
served by ascribing to A a realist understanding of the sentence "J".

Wright prefaces his discussion of Blackburn by referring to other
philosophers - among them John McDowell - who have essayed a similar
line of response to Dummett, and it is clear that Wright intends his

criticisms of Blackburn to have more general application. This intention
might seem rather ambitious if extended to McDowell, however, given that
the latter's version of the argument not only offers an account of the
mechanism whereby speakers might manifest their knowledge of realist
truth conditions, but also includes a specific rebuttal of strategies
employed by Wright - in earlier papers - and Dummett to counter the
realist appeal to such a mechanism. We must therefore look more closely
at McDowell's paper to determine if he is able to succeed where Black-
burn might be thought to have failed.

McDowell, like Blackburn, charges that the Manifestation argument
rests upon the assumption that only certain contexts of language use can
properly be cited by the realist who seeks to demonstrate that a know-
ledge of R-Truth conditions can be manifested in linguistic practice.
He claims that Dummett's argument

> turns on the assumption that, when we seek to ground
> a theory of truth-conditions in linguistic practice,
> the behaviour to which we may look is restricted to
> (behaviour construable as) acknowledgements of the
> truth or falsity of sentences, in response to the recog-
> nizable obtaining, or not, of those truth-conditions. 114

He claims that this assumption rests upon a further premiss concerning
"the character of the internal articulation in a theory of meaning":
more specifically, he attributes to Dummett the thesis that the artic-
ulation of a theory of meaning into a theory of "sense" and a theory of
"force"

> must match an articulation in the psychological state -
> understanding of the language - which the theory aims
> to characterise, in the following sense: the core theory
> [that is, the theory of sense] must correspond to a com-
> ponent psychological state which is in principle isolable

even the sort of theory of understanding that Dummett himself proposes -
one which eschews an appeal to speakers' "perceptual knowledge" of
what others are saying, and seeks to ground our knowledge of what others
are saying in "our ability to discern patterns in one another's behav-
iour" - cannot meet the requirements of an epistemology of understanding
if we adhere to the severe dictates of such a general epistemological
principle:

> We are confronted by questions such as this: supposing
> our awareness of someone's linguistic behaviour to date
> has run smoothly and without puzzlement, so that the
> behaviour has seemed to conform to a familiar pattern,
> how do we know that this was not a short stretch of a
> pattern which equally fits the behaviour so far, but
> which, over a long stretch, we should find unintell-
> igible? 135

What are we to make of McDowell's claim that "the epistemological
root of anti-realism...is a demand for solid foundations for the know-
ledge involved in understanding utterances," a demand which is "inade-
quately justified in general epistemology, and anyway incapable of sat-
isfaction"?[136] Since he offers almost no textual support for his inter-
pretation of the Dummettian project, the case for this interpretation
must consist in its explanatory virtues - its capacity to make sense
of what would otherwise, in McDowell's opinion, be inexplicable lapses
in philosophical judgment on Dummett's behalf. Accordingly, in presen-
ting the case against McDowell's interpretation, I shall argue that it
lacks such explanatory virtues - not because it fails to provide a
rationale for the supposedly puzzling features of the anti-realist
programme (although I find some of the links in the archeological chain
a little rusty), but because there is an alternative interpretation that

> - one which could in principle be possessed indepen-
> dently of the component psychological state to which the
> supplementary theory [that is, the theory of force]
> corresponds. Isolating the core psychological state,
> given that we are not to conceive it psychologistically,
> requires us to find a behavioural manifestation for it
> in particular, separable from what we might go on to
> count as a behavioural manifestation of the psycholog-
> ical state corresponding to the theory of force.   115

As we shall see below, McDowell's archeological explorations of

Dummett's thought proceed to still deeper levels, terminating in what

he terms "the deep epistemological roots of anti-realism."[116] Before

engaging these matters, however, we must examine McDowell's suggestions

as to how the semantic realist can satisfy the "manifestation" require-

ment once Dummett's restrictions on manifestation are removed. He claims

that we can justify semantic realism by requiring that a sub-theory of

the theory of meaning specify "the contents of...assertions which

could be made by uttering the language's indicative sentences."[117] We

may then regard the theorems of this sub-theory as statements of what

are in fact truth conditions for the mentioned sentences:

> There is a truistic connection between the notion of the
> content of an assertion and a familiar notion of truth
> (one whose significance we might think of as fully fixed
> precisely by this connection); the connection guarantees,
> as the merest platitude, that a correct specification of
> what can be asserted by the assertoric utterance of
> a sentence cannot but be a specification of a condit-
> ion under which the sentence is true.   118

There is no requirement, in a theory of meaning of the sort envisaged,

that there be some further sub-theory in which the truth conditions of

sentences are specified independently of the assertoric use to which

they can be put, and thus no requirement that there be some manifes-

tation of a knowledge of those truth conditions distinct from the behav-

iour, whatever it be, that manifests speakers' knowledge of the theorems
of the sub-theory specifying the contents of assertions:

> A specification of a truth-condition is not something
> from which, in conjunction with suitable statements
> from a substantial theory of force, a specification
> of the content of assertions which one might effect by
> uttering the sentence in question is to be derived...
> The specifications of truth-conditions in which we are
> interested simply are specifications of the contents
> of assertions, and that is why we are interested in
> them. What entitles them to figure in the theory of
> meaning is whatever entitles us to suppose that they
> correctly determine what people would be asserting
> in assertoric utterances of the sentences in question.   119

McDowell contends that the foregoing strategy produces a theory
of meaning that both (a) satisfies the "manifestation" requirement,
and (b) is realist in nature. "Realism", for McDowell, is equivalent
to what we have termed "semantic realism": it is "the thesis that a
theory of meaning for a language can give a central role to the notion
of conditions under which sentences are true, conceived as conditions
which we are not, in general, capable of putting ourselves in a position
to recognise whenever they obtain."120 He argues as follows:

> Knowledge of what a sentence can be used to assert is
> knowledge which can be directly manifested, on appro-
> priate occasions, by using the sentence in such a way as
> manifestly to assert precisely that...Specifications of
> contents of potential assertions are, by way of our
> platitude, specifications of conditions under which
> the sentences used to effect those assertions would be
> true. Now if a sentence lacks an effective decision
> procedure. then the condition which any competent speaker
> knows that he would be asserting to obtain if he used
> the sentence in order to make an assertion...is ex hypo-
> thesi not a condition whose obtaining, if it does obtain,
> a competent speaker can be sure of being able to put him-
> self in a position to recognise. Thus...we seem to have
> equipped ourselves with a kind of realism: a description
> of linguistic competence which makes central use of the
> idea that speakers have a knowledge of conditions which

> they are not, in general, capable of recognising whenever
> they obtain.   121

In other words, once we recognise the truistic connection between the
assertoric content of an utterance and the truth conditions of the sen-
tence used to make that utterance, semantic realism follows given only
the further premisses that (1) a speaker's using "S" to assert that p
manifests knowledge that "S" can be used to assert that p; (2) we do
use sentences to assert propositions; and (3) some of the sentences that
are so used by us are "in principle undecidable", e.g. "John is in
pain."122

It should be clear, given our earlier discussion of Strawson, how
Dummett and Wright would respond to this line of reasoning. They would
maintain that the "realism" established by McDowell's argument is not
the kind of realism at issue in the debate over the proper form for
the theory of meaning. For the crux of McDowell's argument seems to be
the claim that speakers can manifest knowledge of what is asserted by
"in principle undecidable" sentences. But, we may recall, the burden of
the Dummett/Wright response to Strawson is that the anti-realist can
grant this point without prejudice - for, so they argue, this claim
entails semantic realism only if we also grant Strawson's implicit
contention that a speaker cannot grasp what is asserted by the assertoric
use of a sentence unless she knows an R-Truth condition for that sentence.
In a footnote, McDowell accuses Dummett of being deaf to the point of
Strawson's argument, that point being that "an account which leaves out
truth-conditions eo ipso leaves out what is certainly important, namely
what the speaker would be saying."123. But, as we have seen, neither
Dummett nor Wright misses this point - they simply dispute it, arguing

that the anti-realist can take the "Wittgensteinian" line and (a)
endorse our ordinary talk of "in principle undecidable" sentences hav-
ing truth conditions while (b) offering an account of what it is for
a speaker to grasp those truth conditions, and thereby to know what is
stated by the assertoric use of the sentences, that traffics only in
anti-realistically acceptable notions. McDowell's treatment of the
issues, it might be claimed, is superficial in that it offers, under
the banner of "semantic realism", an account that can be accomodated
by both parties in the semantic realism debate, while failing to broach
the significant question upon which the two parties are divided -
that is, in what does a speaker's knowledge of the assertoric content
of a sentence consist? What grounds are there for ascribing to language-
users a realist understanding of the sentences of the language? It seems
that McDowell is open to the same argument that Wright employs against
Blackburn: What possible explanatory function, in a theory of meaning,
could be performed by such an ascription? We may be tempted to dismiss
McDowell's "realism" as no more than "an expression of grammar", to
cite again Wright's characterisation of the matter.

McDowell, however, anticipates the spirit, and even to some extent
the letter, of this response, and counters that one should not reject
his account as "superficial" unless there is reason to think that a
deeper analysis is available - which, so he argues, there is not.
McDowell's arguments to this effect are not the easiest to follow,.
and they resist summary in the space available. The following account,
however, which slightly restructures the original presentation, may
be sufficiently faithful and sufficiently detailed to both convey the
general thrust of McDowell's reasoning and reveal what I take to be

the crucial misrepresentations of the anti-realist programme upon which that reasoning depends.

An anti-realist charge that McDowell's argument for semantic realism is "superficial" would, he acknowledges, be supported by reference to his failure to explicate such specifically linguistic functions as assertion. More specifically, as we have seen, Dummett might argue that we cannot take as a primitive, requiring no further explanation in an adequate theory of meaning, a speaker's knowledge that a sentence S is used to assert that p - that is, a speaker's knowing that S has a specific assertoric content. McDowell alludes to certain arguments that Dummett would offer in support of this contention: their general import is that a properly fundamental account of any "pattern of rational activity" or human practice should not presuppose, for its intelligibility, a prior understanding of some similar activity.[124] This is obviously not to be read as the claim that a "properly fundamental account" of linguistic practice should be intelligible to one who doesn't have a practical grasp of the very language in which the account is formulated, which would be absurd. The claim, rather, must be that such an account should not make an explicit appeal, in its analysis of the behaviour or practice in question, to a prior understanding of the nature of the practice under scrutiny, or of an analogous practice.[125]

McDowell claims that to impose such a restriction on a semantic theory - to reject as "superficial" any theory that fails to respect this restriction - is to assume that we can and should avail ourselves of what he terms "the cosmic exile's perspective": it is to assume that

any intelligible human activity can be described, in such a way as to reveal its point or significance, from the

perspective of a cosmic exile - a perspective, that is,
which is not to any extent coloured or affected by the
occupant's own involvement in a form of life; for the
capacity of the description to make the activity com-
prehensible is not to depend on any such involvement.   126

He claims that this assumption, as a general thesis, lacks plausibility:

From the perspective of the cosmic exile, there is no
reason to suppose that any point or significance, of
the sort which human activities have, would be discer-
nible in anything...What seems plausible is, rather,
this: if we insist on eliminating dependence upon prior
involvement in forms of life, then we eliminate the
very possibility of understanding.   127

More specifically, he urges that we draw the following conclusions

concerning the Dummettian project of providing a reflective under-

standing of our own linguistic practice:

The possibility of achieving such reflective understand-
ing would essentially depend on a not particularly reflec-
tive mastery of the practice itself. The fact is that,
in consequence of certain training, we find ourselves
equipped with the perspective, on a certain set of lin-
guistic practices, of a comprehending participant.
According to Dummett's restriction, it should be possible
for us to reconstruct that participant perspective in
terms of materials wholly available to the cosmic exile.
According to the counter-thesis, the sense which, from
our participant perspective, we see in our linguistic
behaviour would be invisible from the cosmic exile's
perspective. The counter-thesis permits us to see a
profound truth in this claim: in theorizing about the
relation of our language to the world, we must start in
the middle, already equipped with command of a langu-
age; we cannot refrain from exploiting that prior equip-
ment, in thinking about the practice, without losing
our hold on the sense which the practice makes.   128

In the final paragraphs of his paper, McDowell offers a similar

diagnosis of Wright's contention that the "realism" defended by Mc-

Dowell - which the latter abbreviates as "R" - is "a mere reflection

of grammar". Again, he contends, this criticism presupposes the poss-
ibility of providing a "deeper" analysis, and this, in turn, presupp-
oses that we can achieve the perspective of the "cosmic exile":

> The idea must be this: the conception of a possibly
> verification-transcendent world appears, as it were
> on the surface, in an account of how things strike some-
> one who is uncritically immersed in a language with the
> right sort of "grammar"; but a more critical perspective
> on that linguistic practice is available, from which
> the world, as realistically conceived, drops out of the
> picture...Wright is evidently impressed by this thought:
> if R is a matter of "grammar", then the possibly veri-
> fication-transcendent world which R countenances is in
> some sense a reflection of our ways of talking about
> it. But if it is true that we cannot think about the
> relation between language and reality except from the
> midst of language as a going concern, then we must in-
> sist on this: if the "reflection" thesis is a truth,
> then it is a transcendental truth, the sort of thing
> that shows but cannot be said. For there is no stand-
> point from which we can give a sense-making character-
> isation of linguistic practice other than that of immer-
> sion in the practice; and from that standpoint our
> possibly verification-transcendent world is certainly
> in the picture. If the "reflection" thesis licenses an
> anti-realism, then it is a transcendental anti-realism,
> one which need not clash with a conviction of the iner-
> adicable necessity of R in our making sense of our-
> selves. [129]

I shall return below to a critical assessment of McDowell's claim
that the sort of meaning-theoretic anti-realism defended by Dummett
and Wright requires the "cosmic exile's perspective", and is therefore,
at best, a "transcendental" thesis, given that our perspective, in our
attempts to achieve a reflective understanding of our linguistic prac-
tice, is never that of the "cosmic exile". Prior to this, however, I
shall critically examine McDowell's further contention that Dummett's
(purported) commitment to the "cosmic exile's perspective" is rooted
in more fundamental features of the Dummettian programme, and, more

specifically, in concerns about the "epistemology of understanding".
I shall argue that this contention reveals a misconception of the anti-
realist programme - a misconception which also underlies McDowell's
claim that we cannot hope for a philosophical account of linguistic
practice that is any "deeper" than his own account R.

McDowell's claim, as we have seen, is that Dummett refuses to
allow the notion of "knowing that S asserts that p" to be a primitive
in the theory of meaning because he believes that the only notions
that are acceptable as primitives in such a theory are ones that would
be available to the "cosmic exile". If McDowell's claim is correct, an
obvious question is why Dummett should hold to the latter thesis.
McDowell's answer to this question is that Dummett "thinks epistemolog-
ical considerations make it compulsory."[130] In fact, the opening remarks
of McDowell's paper make it clear that he believes that the underlying
motivation for Dummettian anti-realism as a whole is epistemological
in nature:

> Fellow-speakers of a language normally understand one
> another's utterances; that is to say, they know what the
> utterances mean. We cannot, then, be content, with a con-
> ception of meaning which makes it a mystery how those
> states can be knowledge, and consequently how there
> can be such a thing as understanding a language. The
> point of the slogan, "meaning cannot transcend use",
> is to suggest a constraint imposed on our conception
> of meaning by the requirements of a credible epistemol-
> ogy of understanding: that the significance of utter-
> ances in a language must, in general, lie open to view,
> in publicly available facts about linguistic behaviour
> in its circumstances. [131]

McDowell contends that the "manifestation" requirement, as an express-
ion of the aforementioned "constraint", is motivated by epistemologic-
al considerations, and that Dummett's reason for rejecting semantic

realism is that it poses unacceptable epistemological problems. This comes out more clearly in subsequent discussion of the (purportedly) epistemological foundations of Dummett's (purported) attraction to the perspective of the "cosmic exile". What Dummett wishes to explain, it is claimed, is how a person who understands the utterances of another can know what the other is saying. McDowell's explanation of this phenomenon consists in an appeal to

> facts which are overtly available (so that conviction
> that they obtain need not be a matter of speculation
> as to something hidden behind what is overtly avail-
> able), but awareness of which is an exercise of a percep-
> tual capacity which is not necessarily universally
> shared. Command of a language is partly constituted by
> just such a perceptual capacity; one whose acquisition
> makes a new range of facts, not hitherto within one's
> perceptual ken, available to one's awareness.   132

He maintains that a speaker's knowledge of what other speakers are saying is perceptual knowledge, available to those who have mastered the language in question. He further claims that Dummett's reason for rejecting this sort of solution to the problem about the epistemology of understanding is also the reason for his attraction to the perspective of the "cosmic exile":

> Dummett's restrictive principle - his insistence on the
> cosmic exile's perspective - would be mandatory if, and
> I think only if, an adequate answer to [the question how
> speakers can know what other speakers are saying] were
> required to take this form: the state's right to the
> title of knowledge must be demonstrated by displaying
> that it consists in awareness of some fact whose avail-
> ability to the knower we can assume without begging
> the question; that is, without prejudging the status of
> our attribution of knowledge.   133

In other words, according to McDowell, Dummett requires that an adequate

solution to the epistemological problem explain A's knowing what B is
saying in terms of (a) A's awareness of some fact, where (b) that
awareness does not itself consist in A's knowing (e.g. perceptually)
what B is saying. But, since the effect of this requirement is to pro-
scribe an appeal, in a theory of meaning constrained by the demands
of an adequate epistemology of understanding, to capacities which A
may acquire in virtue of learning the language spoken by B, it amounts
to the requirement that the theory of meaning restrict itself, in its
primitive vocabulary, to notions that would be available to the "cosmic
exile".

. We arrive, at last, at the final step in McDowell's archeology
of Dummettian anti-realism. We ask, why does Dummett impose the above
requirement upon an adequate solution to the problem in the epistemology
of understanding? The answer, according to McDowell, is that Dummett is
presupposing a general epistemological principle of the following sort:
"The title of a state to count as knowledge depends on there being facts,
unproblematically available to its possessor, which constitute a guaran-
tee...that the content of the putative knowledge is true" ("Cartesian
epistemology"); or, if we permit "criterial" justification that falls
short of providing a guarantee, "the right of some state to count as
knowledge turns on the strength of a ground which the knower has for
the truth of what he knows."[134] McDowell maintains that any such general
epistemological principle is unsound, citing, in support of this claim,
familiar "externalist" criticisms of "criterial epistemology" such as
its supposed inability to account for the possibility of acquiring
knowledge on the basis of the testimony of others. He also argues that

even the sort of theory of understanding that Dummett himself proposes -
one which eschews an appeal to speakers' "perceptual knowledge" of
what others are saying, and seeks to ground our knowledge of what others
are saying in "our ability to discern patterns in one another's behav-
iour" - cannot meet the requirements of an epistemology of understanding
if we adhere to the severe dictates of such a general epistemological
principle:

> We are confronted by questions such as this: supposing
> our awareness of someone's linguistic behaviour to date
> has run smoothly and without puzzlement, so that the
> behaviour has seemed to conform to a familiar pattern,
> how do we know that this was not a short stretch of a
> pattern which equally fits the behaviour so far, but
> which, over a long stretch, we should find unintell-
> igible?   135

What are we to make of McDowell's claim that "the epistemological
root of anti-realism...is a demand for solid foundations for the know-
ledge involved in understanding utterances," a demand which is "inade-
quately justified in general epistemology, and anyway incapable of sat-
isfaction"?[136] Since he offers almost no textual support for his inter-
pretation of the Dummettian project, the case _for_ this interpretation
must consist in its explanatory virtues - its capacity to make sense
of what would otherwise, in McDowell's opinion, be inexplicable lapses
in philosophical judgment on Dummett's behalf. Accordingly, in presen-
ting the case _against_ McDowell's interpretation, I shall argue that it
lacks such explanatory virtues - not because it fails to provide a
rationale for the supposedly puzzling features of the anti-realist
programme (although I find some of the links in the archeological chain
a little rusty), but because there is an alternative interpretation that

is (a) equally adept at explaining the data, (b) less extravagant in its attribution of questionable philosophical doctrines, (c) text-ually more plausible as a reading of Dummett, and (d) available to pro-ponents of the anti-realist programme even if McDowell were to be corr-ect in his account of Dummett's motivation.

Let me refer to McDowell's interpretation of Dummett's arguments for semantic anti-realism as the "epistemological" interpretation. The epistemological interpretation is supposed to explain at least the following features of the anti-realist arguments: firstly, the demand that the theory of meaning satisfy the "manifestation" requirement, and, secondly, the restrictions placed upon the sorts of facts to which one may appeal in attempting to satisfy the "manifestation" requirement - the restrictions which McDowell characterises in terms of the perspec-tive of the "cosmic exile". The former, according to the epistemologic-al interpretation, is an expression of the constraints imposed on the theory of meaning by the requirements of an adequate epistemology of understanding: only if we require that the significance of sentences in a language L, as understood by competent speakers of L, be manifestable in the public use of the language, can we hope to explain how speakers can know what other speakers are saying in their assertoric use of the sentences of L. As to the second of the above features of the anti-realist arguments, the epistemological interpretation construes this as a consequence of Dummett's adherence to some sort of "criterial" epistemological principle: only if we restrict ourselves, in our attempts to satisfy the "manifestation" requirement, to facts that would be available to the "cosmic exile" can our grounds for attrib-uting specific content to the utterances of others confer upon our

resultant beliefs the degree of confirmation necessary if those beliefs
are to count as knowledge.

   In earlier sections of this chapter, I have sketched an alternat-
ive interpretation of the features of the Dummettian project under dis-
cussion. Let me term this alternative reading the "methodological"
interpretation. According to the methodological interpretation, both
the "manifestation" requirement and the restrictions on the satisfaction
of that requirement reflect the nature of our explanatory goals in the
pursuit of a theory of meaning. On this account, the "manifestation"
requirement, and, more generally, the idea that meaning should not
transcend use, is motivated, not by the epistemological concern that
unmanifestable meanings could not be known by other speakers of the
language, but, rather, by the essentially methodological concern that
such meanings, to the extent that they can play no part in the communal
use of language, can have no explanatory function in a theory of mean-
ing that aims to provide a systematic theoretical representation of
that communal use. The restrictions on the satisfaction of the "manifes-
tation" requirement, on the other hand, are not expressions of dubious
epistemological principles, but, rather, reflect the opinion that a
theory of meaning which takes certain central features of linguistic
understanding as primitives, after the fashion of McDowell's R, will
be methodologically flawed in that it fails to provide a full elucid-
ation of the practical abilities that it is the task of a theory of
meaning to explain.

   I have already provided some textual and systematic support for
the methodological interpretation in earlier sections of this chapter -
for example, its capacity to explain why Dummett permits the ascription

to speakers of a knowledge of truth conditions for sentences that are
decidable in principle, and not merely in practice. I shall now attempt
to provide additional support, drawing, in particular, upon those passages that McDowell provides in support of his interpretation. Let me
begin by reiterating Dummett's characterisation of the nature and purpose of a theory of meaning.[137] The philosophical project, we may recall, is not to actually produce such a theory, conceived as "a detailed
specification of the meanings of all the words and sentence-forming
operations of the language, yielding a specification of the meaning of
every expression and sentence of the language."[138] What is sought,
rather, is an enunciation of the general principles for producing such
a theory. The purpose of such an enunciation is to solve philosophical
problems "surrounding the concept of meaning and related notions."
One of these principles links meaning to an understanding of the manner
in which language operates:

> To grasp the meaning of an expression is to understand
> its role in the language: a complete theory of meaning
> for a language is, therefore, a complete theory of how
> the language functions as a language,....that is, of
> how its speakers communicate by means of it: here
> "communicate" has no more precise significance than
> "do whatever may be done by the utterance of one or
> more sentences of the language".   [139]

Notice that, in this passage, Dummett deliberately deflates the
apparent epistemological implications of characterising a theory of
meaning as a theory which explains how speakers communicate. On the
epistemological interpretation of the anti-realist project, the crucial constraint on a theory of meaning is that it be able to explain
how genuine communication is possible - how A can know what B is saying

and thus how B can communicate with A through the medium of language.
By glossing "communication " as whatever speakers do by uttering senten-
ces, however, Dummett defuses any such epistemological implications of
the use of the term in characterising a theory of meaning. Not only is
it rendered completely unproblematic that "communication" actually occ-
urs when sentences are uttered, but also it is left quite open whether
the occurrence of such "communication" either raises or resolves any
epistemological problems. This effective distancing of the project of
constructing a theory of meaning from epistemological concerns is even
clearer, I think, in a passage which McDowell himself cites to illus-
trate Dummett's supposed commitment to a criterial conception of know-
ledge. The passage in question is offered by McDowell as a context in
which Dummett speaks of knowledge as something requiring a "guarantee".
Here, however, is the passage containing the latter reference:

> To suppose that there is an ingredient of meaning which
> transcends the use that is made of that which carries the
> meaning is to suppose that someone might have learned
> all that is directly taught when the language of a-
> mathematical theory is taught to him, and might then
> behave in every way like someone who understood that
> language, and yet not actually understand it, or under-
> stand it only incorrectly. But to suppose this is to make
> meaning ineffable, that is, in principle incommunicable.
> If this is possible, then no one individual ever has a
> guarantee that he is understood by any other individ-
> ual; for all he knows, or can ever know, everyone else
> may attach to his words or to the symbols which he em-
> ploys a meaning quite different from that which he att-
> aches to them. A notion of meaning so private to the in-
> dividual is one that has become completely irrelevant
> to mathematics as it is actually practised, namely, as
> a body of theory on which many individuals are corporately
> engaged, an enquiry within which each can communicate
> his results to others. ·   140

At first glance, it might seem that this passage supports Mc-

Dowell's interpretation of Dummett. For there is, indeed, an apparent
tacit endorsement of the view that knowledge requires a "guarantee",
and it also appears that the motivation underlying the requirement that
meaning not transcend use is epistemological in nature - namely, the
epistemological implications vis-a-vis our ability to understand one
another if we allow meaning to transcend use. Closer inspection, however,
reveals that it would be incorrect to draw such inferences from the pass-
age. In the first place, it is not the case that Dummett assumes, in
the context of a concern with what speakers can properly be said to
know, that knowledge.requires a "guarantee"; rather, in the context of
a concern with those things for which speakers could have a guarantee,
Dummett characterises matters in terms of "what they know". That this
difference is not trivial is shown by the fact that the difficulty
Dummett is raising would not be overcome by the expedient of pointing
out that speakers can know - in the "externalist" sense - that they
are or are not understood by others; for the difficulty is the lack of
subjective certainty, whether or not this be classified as a lack of
"knowledge". Dummett's usage of "know", in this context, accords with
the vernacular. It may, nonetheless, be expressively infelicitous, or
even philosophically irresponsible. But it would surely be even more
irresponsible to attribute a criterial conception of knowledge to
Dummett on the basis of this passage alone, especially if one wished
to ground upon such an attribution a diagnosis of the ills of the entire
anti-realist programme in the philosophy of language.

Secondly, Dummett does not argue, from the epistemological con-
sequences of allowing meaning to transcend use, to the conclusion that
the theory of meaning must not permit meaning to transcend use, or,

more specifically, to the conclusion that an adequate theory of meaning must satisfy the "manifestation" requirement. Rather, he points to the epistemological implications of manifestation-transcendent meanings, and then argues that such "meanings" are "completely irrelevant" to actual linguistic practice in mathematics, and , by extension, in other areas of discourse. But a theory of meaning, as we have seen, is an account of "how the language functions as a language", and its explanandum is our actual linguistic practice. If, then, manifestation-transcendent meanings are "completely irrelevant" to our linguistic practice, then they are also presumably completely irrelevant to an explanation of that practice within the context of a theory of meaning - unless, of course, we adopt the third realist strategy, to be examined in the following section, and permit meaning to transcend what is understood by the competent speaker. The justification for the "manifestation" requirement, therefore, is neither that manifestation-transcendent meanings pose epistemological difficulties, nor that we can somehow show that there are no such meanings, but, rather, that, even if there are such meanings, they should be excluded from a philosophical theory of meaning on methodological grounds - roughly, the Occamian principle that an analysis should eschew quantifying over entities that make no contribution to that analysis. Thus the passage cited above supports the methodological interpretation, and not the epistemological inter-pretation upon whose behalf it is summoned by McDowell.

Similar considerations apply, I think, to the only other textual evidence which McDowell adduces in support of his interpretation. To be strictly accurate, the material in question is not so much offered as evidence, as incorporated into the presentation of Dummett's views

in such a manner as to lend credence to the idea that Dummett's concerns, in challenging semantic realism, are primarily epistemological. McDowell alludes to Dummett's hostility to psychologism in the philosophy of language, a position which the former characterises as "a conception according to which the significance of others' utterances is a subject for guesswork or speculation as to how things are in a private sphere concealed behind their behaviour."[141] He later suggests that the reason why Dummett imposes a restricted version of the "manifestation" requirement upon the theory of meaning is that he takes this to be the only alternative to psychologism, and that he rejects psychologism because it fails to satisfy the epistemological constraints imposed by the epistemology of understanding. Dummett, he claims, believes that we are confronted with a dilemma:

> Either the acquisition of knowledge in which, on appropriate occasions, the understanding of a language is exercised consists in awareness of [facts detectable from the perspective of the "cosmic exile"]; or else we must be conceiving it psychologistically, as involving hypotheses about inner states of the speaker lying behind the behaviour. Avoiding psychologism, Dummett opts for the first horn.   142

But, as we have just seen, there is independent textual evidence against the suggestion, here, that Dummett argues from the epistemological consequences of a view - such as psychologism - that traffics in manifestation-transcendent meanings, to the need to impose the "manifestation" requirement, restricted or otherwise, upon the theory of meaning.

It is perhaps worth clarifying just what bearing Dummett's hostility to psychologism does have upon his conception of a theory of meaning, given what I believe to be a misconception of this matter in a

footnote which McDowell appends to his explanation, in terms of the epistemological interpretation, of Dummett's supposed attraction to the perspective of the "cosmic exile":

> I used to think it an adequate diagnosis to ascribe to Dummett the idea that a theory of meaning for a language is an object of implicit knowledge of speakers, whereby they guide their linguistic behaviour...But the diagnosis is at best incomplete; for the idea of inward consultation of a recipe for correct speech is profoundly psychologistic, whereas Dummett is hostile to psychologism, and disavows it in precisely those passages in which he offers the conception of a theory of meaning as an object of implicit knowledge...What I am going to suggest is, in effect, this: Dummett is attracted by that psychologistic idea (and blinded to its psychologistic character) because, ironically, he thinks it is the inexorable outcome of considerations whose point lies in opposition to psychologism. 143

If one consults the passages cited by McDowell in support of this rather unflattering diagnosis[144], however, one might conclude that it is the critic rather than his subject whose perception of matters is flawed. For he misconstrues both what Dummett is saying about "implicit knowledge" and what he is saying about "psychologism".

Firstly, Dummett's point about a theory of meaning being "an object of implicit knowledge" is the familiar one that, in constructing a theory of meaning which gives a theoretical representation of the practical abilities of competent speakers of a language - representing those abilities in terms of knowledge of a set of propositions - we cannot in general require that the knowledge ascribed to speakers be something that can be _explicitly_ manifested in the ability to state the appropriate theorems of the theory of meaning. Rather, for a large segment of the language, the speaker's knowledge must be construed as _implicit_ knowledge. But, we will recall, Dummett's central claim is that

the ascription of implicit knowledge, in the construction of a theory of meaning, is to be strictly governed by the "manifestation"-requirement. The implicit knowledge ascribed to speakers by a theory of meaning that conforms to this requirement cannot, therefore, be a knowledge of manifestation-transcendent meanings, and, as a consequence, the ascription of such knowledge is not attended by any of the epistemological consequences of psychologism noted by Dummett in the passage cited earlier.

Secondly, Dummett's "hostility", in the passages cited by McDowell, is directed towards the idea that a theory of meaning is "intended as a psychological hypothesis", that its concern is to "describe any inner psychological mechanisms" that might explain a speaker's having any specific practical linguistic abilities.[145] But, as was stressed in the discussion of Devitt in subsection 1 above, to deny that a theory of meaning is a "psychological hypothesis" in this sense is not to deny what I have termed "token psychologism", and there is no reason for Dummett to be hostile to the latter position - indeed, there are good reasons to endorse it, if the theory of meaning is to be accorded some sort of "explanatory" status. We might put the point as follows: what Dummett rejects is the idea that a theory of meaning should have anything to do with private meanings of public language (where "privacy" is taken to entail manifestation-transcendence); but to reject this is not to reject the idea of a private grasp of public meanings.

We should note, finally, that even if the epistemological interpretation were a correct reading of Dummett, there is no reason why the aspiring anti-realist should not defend the restricted version of the "manifestation" requirement in the manner of the methodological inter-

pretation. Wright's response to Blackburn provides an example of an anti-realism of this kind. He explicitly rejects the idea that the problem with semantic realism is an inability to answer the skeptic. He can therefore cheerfully accept McDowell's charge that the anti-realist is not able to answer the skeptic either. As we have seen, Wright explicitly endorses what I have termed the methodological interpretation of the anti-realist programme.

Let me now return to McDowell's contention that the semantic anti-realist's claim to offer something "deeper" than R presupposes the availability of the perspective of the "cosmic exile", and that, as a consequence, the proposed anti-realism is at best a transcendental truth, incapable of being stated. I shall suggest that such a characterisation of the anti-realist programme is indeed plausible if one subscribes to the epistemological interpretation of the programme, but that it is most implausible if one subscribes - as I have suggested one should - to the methodological interpretation.

The supposedly transcendental nature of semantic anti-realism - which McDowell, unhelpfully for our purposes, terms "meaning-theoretic anti-realism" - is affirmed by McDowell in the following remarks at the end of his paper:

> The transcendental realist claims that from the cosmic
> exile's perspective, one would be able to discern rela-
> tions between our language and a realistically conceived
> world. Anti-realists justifiably recoil, but in differ-
> ent ways. The meaning-theoretical anti-realist recoils
> into giving a different picture of how things would look
> from that perspective; but the right course is to set
> our faces against the idea of the cosmic exile.   146

The contention is that the "more critical perspective" on our linguis-

tic practice that is required if we are to offer something "deeper" than R is available only from the perspective of the "cosmic exile", where one is able to "discern relations between our language and a realistically conceived world." The argument for semantic anti-realism, on this reading, is that, from such a perspective, one can discern that there are no such relations conferring realist content on the sentences of our language - that, in McDowell's characterisation of Wright's position, "the possibly verification-transcendent world which R countenances is in some sense a reflection of our ways of talking about it."[147]

Suppose, for the moment, that the motivation underlying the anti-realist programme in the philosophy of language really were epistemological, as McDowell maintains. Then to restrict the meanings ascribable to the sentences of our language to their "manifestable" meanings would not, by itself, be sufficient to assuage our epistemological ills as putative anti-realists: for we would also require an assurance that such a restriction captures the meanings that the sentences uttered by our fellow-speakers actually possess. Otherwise, we would have no reason to believe that, from the fact that our anti-realist theory of meaning tells us that we know what others are saying, we can infer that we really do understand what others are saying. But the only apparent way of achieving such assurance requires that we discern, from the same vantage-point that the "transcendental realist" claims to occupy, that those relations asserted to obtain by the realist - relations capable of conferring manifestation-transcendent content on our language - do not obtain. A similar analysis would apply to Wright's contention that R is "a reflection of grammar". It is therefore plausible to suppose

that an epistemologically-motivated semantic anti-realism would make at least a tacit appeal to McDowell's "perspective of the cosmic exile".

What, however, if we reject the "epistemological" interpretation of the anti-realist programme, and cleave, instead, to the methodological interpretation? Do we still have grounds for thinking that the anti-realist is, at least implicitly committed to the perspective of the "cosmic exile", and that the anti-realist thesis is plausible only if construed transcendentally (and, thereby, construed as unstatable, if we follow-Wittgenstein on this)? Let me briefly suggest why we should conclude that we do not have grounds for ascribing such a commitment to the methodologically-motivated anti-realist.

In the first place, the methodological anti-realist programme in the philosophy of language can endorse the "manifestation" requirement, as a principle for excluding manifestation-transcendent meanings from the analysis of linguistic practice provided by a theory of meaning, without having to establish that <u>there are no such meanings</u>. All that matters is that, to the extent that some element of meaning really is "manifestation-transcendent", it can play no part in the public linguistic practice that a theory of meaning seeks to elucidate. Thus the methodological anti-realist has no need to appeal to the perspective of the "cosmic exile" to establish the non-existence of such manifestation-transcendent meanings. Furthermore, the anti-realist's concern to exclude, from the theoretical representation of the linguistic abilities of competent speakers, anything that is not strictly accountable to the actual or possible employment of the language seems to preclude ascribing a transcendental status to the anti-realist theory of meaning itself.

But perhaps we are missing McDowell's point, here, for it is not in the "manifestation" requirement itself, but in the restrictions imposed on the satisfaction of that requirement, that McDowell perceives the traces of the perspective of the "cosmic exile". The issue, we may recall, is whether it is "superficial" to take a notion such as "knowing that S is used to assert that p" as a primitive in a theory of meaning. McDowell's argument, which I shall repeat for convenience of exposition, is that the sort of "reflective understanding of our own linguistic practice" that we seek in a theory of meaning may presuppose that we possess "a not particularly reflective mastery of the practice itself":

> The fact is that, in consequence of certain training, we find ourselves equipped with the perspective, on a certain set of linguistic practices, of a comprehending participant. According to Dummett's restriction, it should be possible for us to reconstruct that participant perspective in terms of materials wholly available to the cosmic exile. According to the counter-thesis, the sense which, from our participant perspective, we see in our linguistic behaviour would be invisible from the cosmic exile's perspective. The counter-thesis permits us to see a profound truth in this claim: in theorising about the relation of our language to the world, we must start in the middle, already equipped with command of a language; we cannot refrain from exploiting that prior equipment, in thinking about the practice, without losing our hold on the sense which the practice makes.   148

In order to evaluate this argument, we need to clarify two matters. Firstly, on what grounds might the anti-realist argue that, if the theory of meaning is to provide us with the sort of "reflective understanding of our own linguistic practice" that we seek, then it cannot take such notions as "knowing that S is used to assert that p" as primitive. Secondly, to what extent does McDowell's claim, that understanding our lin-

guistic practice - grasping its "sense" - presupposes the perspective of "a comprehending participant", undermine the possibility of achieving the sort of meaning-theory that the anti-realist seeks? If we are to answer the first of these questions, I think we need to ask a further question, namely, what sort of "reflective understanding of our own linguistic practice" can we hope to gain from a theory of meaning, as part of a more general philosophy of language? Or, given Dummett's claim that a completed theory of meaning would provide "a complete theory of how the language functions as a language", what insight into these sorts of matters might we hope to derive from philosophy?

I think the answer to the latter questions lies along the following lines. What we are trying to do, in enunciating the general principles that would govern the construction of a theory of meaning for an entire language, is to provide the conceptual foundations for what might be termed "intensional discourse". By "intensional discourse", I understand our practice of describing language-using creatures like ourselves as being in intensionally-individuated mental states; as producing utterances having propositional content; and - if we follow Putnam in our account of what have traditionally been taken to be "extensional" features of language - as referring to particular entities by means of particular linguistic items. Intensional discourse, so construed, will include such familiar actiyities as attributing beliefs and desires to people, taking people to understand or misunderstand what we are saying, taking ourselves to understand or misunderstand what others are saying, translating between different languages, and deciding whether earlier generations of inquirers were really talking about the same things as us. Philosophical problems about meaning are

a subset of the conceptual problems generated by such practices, and our goal, in our essays at a theory of meaning (and in philosophy of language in general), is to provide a conceptual framework that will help us to make sense of those practices. It is this sort of "sense" that we hope to glean through arriving at "a reflective understanding of our linguistic practice". And the importance of arriving at such an understanding resides in the importance of intensional discourse itself, as what, following Putnam's discussion in his recent Carus Lectures, we may term as "indispensable version" - a version that plays an indispensable role in our interactions with and understanding of our fellow human creatures.

It should be apparent that, to the extent that we view the philosophical pursuit of a theory of meaning in this way, we will regard as "superficial" an account such as McDowell's R. For such an account, in taking our knowledge of what a sentence asserts as primitive, fails to provide an elucidation of a central feature of our intensional discourse, and leaves us without a theoretical representation in terms of which to make "sense" of the relevant part of our practice. But is McDowell correct in claiming that the quest for anything "deeper" that R is chimerical because we can only theorise about language if "already equipped with command of a language", and because the "sense" of our linguistic practices can be grasped only by one who possesses "the perspective of a comprehending participant"? What do the two reasons for not seeking anything "deeper" than R - cited in the previous sentence - amount to?

The first reason might be read as the claim that we cannot make sense of a theoretical representation of a practical ability unless we

already understand the language in which the representation is formulated. This, however, is quite uncontroversial, as I noted earlier. If there is a deeper point here, then it must reside in the sort of difficulty specified in the second of the above reasons. The claim, here, seems to be that one cannot really understand a theoretical analysis of a practical ability unless one has experience of exercising such an ability (or, perhaps, of exercising a similar ability). But this claim seems highly suspect, and seems to rest upon a confusion of two senses of "sense". An example may be helpful here. An individual who has never played American Football may be unable to "see the sense" in the exercise of those skills that constitute the playing of the game, where failing to "see the sense" in an activity is a matter of failing to understand why people would engage in that form of activity as a whole. But the same individual might be perfectly able to "make sense" of American Football - that is, understand what is going on within the game - because he has mastered a theoretical representation of the game - i.e. a primer of rules and strategies. But we do not expect such a theoretical representation of a form of activity to enable one who masters it to "see the sense" of that activity, in the former sense of "sense" - although mastering the internal structure of an activity may help one to understand what people "see in it". Now, it is surely in only the first sense of "sense" that it might plausibly be held, of at least some activities, that one cannot "grasp the sense" of the activity unless one attains the perspective of "a comprehending participant". But it is surely in the second sense of "sense" that we look to a theory of meaning, as a theoretical representation of a practical ability, to help us "make sense" of our linguistic activities.

It is perhaps unfair to accuse McDowell of such an elementary
mistake as the conflation of these two senses of "sense", although he
lends himself to such a charge by arguing that a theoretical represen-
tation of our linguistic abilities from the perspective of the "cosmic
exile" would be inadequate because it would fail to capture "the sense
which, from our participant perspective, we see in our linguistic
behaviour," and would involve "losing our hold on the sense which the
practice makes". But I am unable to come up with a more plausible ver-
sion of his argument, and, in the absence of such, I think we must
conclude that he has given us no reason to doubt that the "deeper"
analysis of our linguistic practice pursued by the anti-realist is
obtainable.

4) Content and Understanding

We now come to the third and final strategy adopted by critics of
Dummett's arguments for semantic anti-realism - the claim that, even
if we grant the case against meaning-theoretic realism, this fails to
refute semantic realism because the sentences that we understand may
still _possess_ R-Truth conditions and this fact may be crucial to an
understanding of those very linguistic practices that a philosophical
account of language seeks to elucidate. It is, of course, necessary
for the realist to defend _both_ of these points: for to establish only
the first point - that the sentences that we understand _possess_ R-
Truth conditions - would leave unanswered the challenge, issued at the
beginning of this chapter, that the realist programme in the philosophy
of language is bankrupt. We have already critically examined one variant
on this third strategy, namely, Putnam's "Reference and Understanding"
proposal to combine a verificationist account of linguistic understand-
ing with a realist explanation of the contribution that language makes
to the achievement of our cognitive and non-cognitive goals. I shall
shortly rehearse our critical judgment on this proposal. Prior to that,
however, I shall sketch two further variations on the third strategy
contained in Devitt's already cited paper and in a paper by Colin Mc-
Ginn. I shall focus upon McGinn's presentation, incorporating refer-
ences to Devitt where appropriate.

McGinn argues[149] that Dummett's case against semantic realism
rests upon an optional, and ultimately unacceptable, conception of
" how content is conferred on sentences". This conception, which he
labels the "_dispositional_ conception of content", is said to underlie
the Manifestation argument. It consists in "the idea that all that can

warrant and determine content-ascriptions are the dispositions of speak-
ers in respect of evidence."[150] The dispositional conception of content
seems to be equivalent to a principle that Devitt ascribes to Dummett
in his reconstruction of the Manifestation argument - the principle that
a sentence "has no truth conditions other than those associated with it
recognitionally by the speaker's understanding."[151] McGinn maintains
that this conception of content correctly characterises a constraint
on something akin to the Fregean "sense" of a sentence, and that, to
the extent that a speaker's understanding of a language consists in
a grasp of such Fregean senses, or is properly so represented, the
Manifestation argument succeeds in establishing what we have termed
meaning-theoretic anti-realism. But, he claims, we should reject the
dispositional conception as an _exhaustive_ conception of the content
properly ascribable to the sentences of a language:

> Sense should be explained in terms of use, but ..it is
> false that content is exhausted by sense, so understood:
> the referential truth-conditions of sentences are inde-
> pendently determined and constituted...
> Epistemological sense will consist in assertibility
> conditions and grasp of these will be manifested in much
> the way Dummett suggests, i.e. in the exercise of recog-
> nitional abilities; but reference will supervene on
> facts about how the speaker is embedded in the world -
> upon certain causal or contextual facts perhaps.   152

McGinn further contends that Dummett fails to realise that sen-
tences can have recognition-transcendent content in virtue of the con-
tribution that referential truth conditions make to total content bec-
ause he believes that reference must be determined by sense - or, in
other words, because he holds to a "descriptions" theory of reference.
Once it is acknowledged that reference is a function of "certain

causal or contextual facts...about how the speaker is embedded in the world", we should also recognise that semantic realism is correct, as long as the things with which we causally and contextually interact are what we believe them to be - namely, denizens of a verification- and recognition-independent reality, and _not_, e.g. the "things in the image" that are the referents of terms in vat-English.[153] Devitt takes a similar line in his attack on his version of the "dispositional conception of content". The fundamental disagreement between himself and Dummett, he maintains, is over the relative merits of "descriptivist" and "causal" theories of reference. It is a consequence of the "causal theory", he claims, that "words have referents, and hence sentences have truth-conditions, which are not associated with them recognitionally by the speaker's understanding."[154] (This should have a familiar ring; these are the same sorts of charges that Devitt and Lewis level at the "model-theoretic" argument - and, as we shall see, they are open to the same sorts of responses.)

McGinn addresses directly the question how recognition-transcendent truth conditions that do not enter into an account of understanding a language can nevertheless have a crucial bearing upon the way in which the language functions as a language. He argues that referential truth conditions are required if we are to give an adequate account of the _point_ of language use:

> The point of assertion...is to communicate to one's audience the conditions of the world. A full description of the activity of speech would therefore need to invoke the notion of reference to make sense of what speakers are up to when they assert sentences, but it would be wrong to infer that reference must be _manifest_ in use in the way I have been opposing.   155

He further argues that our actual practices of content-ascription·rest
upon principles that involve a commitment to the sort of recognition-
transcendent content that the realist requires. In support of this claim,
. he appeals to our intuitions concerning certain hypothetical examples; ·
the latter - variants on familiar "Twin Earth" and "brain in a vat"
scenarios - echo Putnam's "pre-critical" arguments against "descriptiv-
·ist" theories of reference. A brief sketch of two of McGinn's examples
will, I think, convey the flavour of his reasoning.

In "Case A", we have the standard "Twin Earth" scenario, where a
term, identified phonetically, has distinct referents in the idiolects
of two linguistic communities, but where, it is assumed, the assertibil-
ity criteria associated with the term in the two communities are not
sufficient to discriminate between the referents. In such a case, McGinn
maintains, one who espouses the dispositional conception of content must
hold that phonetically identical sentences, in the two idiolects,·con-
taining the term in question - say, 'water' - have the same content,
since the linguistic dispositions of speakers in the two communities
with respect to that sentence are the same:

> The same evidence, viz. facts about how liquids look and
> taste, will prompt them to acknowledge the truth of sen-
> tences with different referential content. Speakers on
> earth and twin earth thus acquire the same recognitional
> capacities amd manifest them in the same conditions of
> evidence, but their sentences do not mean the same;
> they have the same assertibility conditions but differ-
> ent truth-conditions.   156

What this purportedly shows is that the dispositional conception of
content is incorrect, for we have at least one instance in which "agree-
ment in use does not entail agreement in content." Nor, according·to

McGinn, is it possible for the anti-realist to evade this conclusion
by either of the following strategies: (a) "Characterising use in vocab-
ulary relating to the environments of the speakers on earth and twin
earth" - this would be unacceptable to an anti-realist such as Dummett
because such externally characterised dispositions would not be reflec-
ted in speaker's knowledge, McGinn maintains; or (b) claiming that
speakers are counter-factually disposed to use the term 'water' differ-
ently, i.e. differently disposed "were they to come to know the scien-
tific nature of the substances" - this fails to solve the problem, McGinn
argues, because speakers may lack the intellectual capacities to arrive
at such a scientific understanding.[157]

   In "Case C", we have what is, for all essential purposes, the same
"brain in a vat" scenario discussed by Putnam, save that it is not all
sentient beings, but only all sentient beings on Twin Earth, that are
permanently isolated brains in a vat. The supposition is that the envatt-
ed brains

> have acquired a language phonetically identical with our
> language of material objects in the same way we have,
> at least in this respect: they have learned under what
> experiential conditions to assent to their sentences.
> In short, they have acquired the same recognitional cap-
> acities in respect of evidence - they know what exper-
> iences justify acceptance of which sentences. They have
> therefore associated the same assertibility conditions
> with their sentences as normally situated people have.   158

As in Case A, McGinn's contention is that we have a difference in con-
tent that cannot be accounted for on the dispositional conception. For,
he argues, there is no difference between ourselves and the vat-dwellers
"in point of recognitional capacities, associated assertibility condit-
ions, and dispositions to use", but " they do differ with respect to

the content of their sentences and their propositional attitudes." The content differs because the referential truth conditions differ, the latter being determined by "the environmental context ....in particular, the causal, epistemic, and action relations in which a person stands to the environment in which he is embedded." In the case of the vat-dwellers, McGinn claims that their sentences "do have truth conditions but... these just consist in their assertibility conditions," since terms in vat-English have only phenomenal referents.[159]

. At this point in his argument, McGinn asks whether the differences in content supposedly illustrated by the above examples is indeed a difference properly characterised in terms of the applicability or non-applicability of a realist semantics. He maintains that this is the proper way to characterise case C: "The difference of content reflects the difference between assigning truth conditions which are irreducible to experiential assertibility conditions and refraining from assigning such conditions, and that difference is the difference between realist and non-realist semantics."[160] To further support this contention, he offers "Case D", where, in the sort of scenario entertained in case A, we have a term, 'S', in both idiolects, that refers to the sensations that Earth bats have when they use the sense of echolocation. It is assumed that Twin Earth bats, while behaviourally and psychologically indistinguishable from Earth bats, lack this sensation. On both Earth and Twin Earth, certain behaviour by bats constitutes the assertibility conditions for S-ascribing sentences. On Earth, McGinn argues, S-ascribing sentences have referential truth conditions distinct from their assertibility conditions - they are true when the sensation S is present

in a given bat's stream of consciousness - but on Twin Earth we must say that such sentences <u>lack</u> truth conditions unless we are behaviourists. Again, he concludes, we have equivalent use attended by difference in content, and the difference in content is one that must be characterised in terms of the applicability of a realist semantics:

> The reductive class consists of statements about behaviour, and the realist claim is that truth for the class of psychological sentences does not reduce to truth for that reductive class - the truth of psychological sentences is verification-transcendent. 161

Note, firstly, that McGinn's attempts, as sketched in the preceding paragraph, to show that his examples bear upon the viability of semantic realism fall prey to the already-cited confusion between semantic realism and ontological realism. Whether or not a given class of sentences has truth conditions that are reducible to "experiential assertibility conditions" bears upon the propriety of adopting an ontological realist stance towards the members of that class. But the <u>semantic</u> anti-realist is not committed to any such reductionist stance. As we saw in our discussion, in subsection 2 above, of Strawson's arguments against anti-realism, the semantic anti-realist has nothing invested in a behaviourist account of the truth conditions of sensation-ascribing statements. Putting aside this confusion, however, we may ask whether McGinn's examples give us any grounds for concluding that we operate with a realist conception of sentence content, and that an adequate philosophical account of our linguistic practice must incorporate such a conception, and, thereby, the notion of R-Truth.

Consider, first, case A. We may recall that McGinn canvasses and rejects two ways in which Dummett might try to bridge the purported gap

between content, as revealed by the example, and "dispositional" content. The second suggestion is that Dummett might appeal to speakers' counter-factual dispositions, conditionalised on the hypothesis that they arrive at a scientific understanding of the nature of the substances that are the referents of the term 'water' in the two idiolects. McGinn dismisses this suggestion on the grounds that speakers may lack the intellectual capacities necessary to fulfil the antecedent of the conditional. In a sense, McGinn is right, but this doesn't support the sort of conclusion that he wishes to draw.

One way in which a linguistic community might lack the intellectual prerequisites for developing a scientific understanding of the nature of different liquids is if the members of that community lacked what might be regarded as a precondition for even attempting to develop such an understanding - namely, the idea of an intelligible order of things distinct from the way in which things present themselves in ordinary sense-experience (or, in more traditional parlance, the idea of an intelligible Being that lies behind the fluctuating world of Becoming). In a less extravagant vein, we can imagine a linguistic community that treats the term 'water' in the manner in which our hypothetical community would presumably treat all referring expressions - as a term that picks out what might be called a "phenomenal kind", a class of things distinguished solely in terms of their capacity to elicit a particular kind of sense-experience in perceivers under specified conditions. Suppose that the linguistic communities in McGinn's case A employ the term 'water' in this way. In such a case, it would not be possible to account for a difference in the content of Earth and Twin Earth utterances of sentences containing the term, 'water', by citing

a difference in the relevant dispositions of members of the two linguistic communities were they to possess a scientific understanding of the chemical differences between the Earth and Twin Earth referents of 'water'. It would not be possible <u>because such an understanding would leave the linguistic practice of both communities unchanged</u> as long as they continued to treat 'water' as a phenomenal kind term. For, as long as the term is so regarded, it refers to whatever possesses the relevant sensible qualities, regardless of chemical composition. And thus the term 'water' will have the same extension in each of the idiolects, and there will be no difference in the content of sentences containing the term for the anti-realist to explain.

Presumably, though, what McGinn has in mind is a situation in which the term 'water' is employed, not as a phenomenal kind term, but as a "natural kind" term. In this case, on the sort of Putnamian "causal" account[162] to which McGinn appeals, members of our two linguistic communities intend, in their use of the term 'water', to pick out a particular kind of "stuff" in the world, instances of which fall under interesting generalisations, and properties of which can feature in explanations of the ways in which the world presents itself to us in experience. Since the "stuff" appropriately related to the employment of the term 'water' in the two linguistic communities is different in nature - being H_2O on Earth and XYZ on Twin Earth - the referential truth conditions of sentences containing the term in the two idiolects are also different, and the sentences will differ in content

But, to the extent that we adhere to the Putnamian analysis of such matters that yields this result, we must also recognise that it is part of our linguistic practice, with respect to sentences contain-

ing a natural kind term such as 'water', that ascriptions of the pred-
icate, "x is water", to things having those sensible qualities that
make up our "stereotype" of water are in principle <u>defeasible</u>. We allow
that two samples of liquid could possess identical sensible qualities
yet fail to be samples of the <u>same</u> liquid. Since the linguistic dispos-
itions of a competent speaker will incorporate a grasp of this feature
of linguistic practice, and since such dispositions entail a counter-
factual disposition to differentiate, in terms of chemical composition,
the referents of 'water' in the two idiolects, there will be a feature
of the use of sentences containing 'water' in the two idiolects that
corresponds to the acknowledged difference in the <u>content</u> of those sen-
tences. Of course, speakers in the two communities <u>ex hypothesi</u> have no
currently available means of distinguishing between the two liquids in
question. But, as we have already seen, anti-realist constraints on
content-ascription are formulated in terms of in principle, and not in
practice, manifestability.

It might be argued, however, that the foregoing analysis fails to
provide an appropriate analogue, in use, for the difference in the con-
tent of sentences containing the word 'water' in the two idiolects. For,
granting that practice in each community incorporates the noted defeas-
ibility condition, the <u>difference</u> in content still stems from the fact
that the things taken as paradigms of water in the two communities are
<u>different substances</u>. This response has, I think, the effect of assimil-
ating case A to case C; so, rather than dispute further concerning the
former, I shall turn to the latter. The crucial contention in both
cases - given this construal of case A - is McGinn's claim that the
anti-realist cannot avail herself of the first of the strategies can-

.vassed in respect of case A - the strategy of "characterising use in vocabulary relating to the/environments of the speakers on earth and twin earth." McGinn contends that this strategy is not available to Dummett because, if speakers' dispositions to use the language are characterised "externally" in this fashion, such dispositions will not be reflected in speakers' knowledge and meaning will no longer be essentially cognitive.[163] The argument, if I understand it correctly, is that (a) "externally" characterised dispositions to use sentences would not be fully manifested in linguistic practice, that (b) we cannot, as a consequence, ascribe to competent speakers knowledge of use, so construed, and that (c) if meaning is identified with externally characterised use, in order to accomodate the difference in content between phonetically identical utterances on Earth and Twin Earth, meaning will transcend what a competent speaker can be said to understand, as given by the theory of understanding.

But this line of reasoning rests upon certain assumptions that we have already had cause to dismiss in another context - the assumptions, central to Hansen's criticisms of the "model-theoretic" argument, that anti-realism presupposes some sort of "Lockean" epistemology, and that an anti-realist conception of linguistic content cannot reflect the (potentially different) ways in which a given way of using a language is embedded in our more general commerce with the world. Consider, here, McGinn's analysis of case C cited above. He claims that there is a difference in content for a given sentence - say "There is a table in front of me", or "T" for short - as uttered by ordinary speakers on Earth and by envatted brains on Twin Earth, a difference that stems from the different "causal, epistemic, and action relations" in which

the two sets of speakers stand to "the environment in which [they are] embedded." But these differences in the environmental embedding of language-use are not, on McGinn's analysis, reflected in "recognitional capacities, associated assertibility conditions, and dispositions to use", for, in all these respects the two communities are supposedly identical: members of both communities have associated with "T" the same "experiential conditions", and have thereby acquired "the same recognitional capacities in respect of evidence". But the latter can be true only if "experiential conditions" are construed as "inner representations", rather than as states of the world that speakers can recognise to obtain. For, if we adopt the latter construal of what "experiential conditions" are, it is clear that members of the two linguistic communities acquire quite different recognitional capacities, and associate the sentence "T" with quite different assertibility conditions. Describing the situation from a "God's Eye" perspective, we will say that Earthlings associate "T" with the presence of a table, whereas Twin Earthlings associate "T" with the occurrence of a "table in the image". Furthermore, this difference will be manifested in the linguistic practice of members of the two communities: Earthlings will manifest their knowledge that "T" is assertible in the presence of tables, and Twin Earthlings will manifest their knowledge that "T" is assertible upon the occurrence of "table in the image" experiences. What neither community will manifest, of course, is an ability to discriminate between the obtaining of the Earth and the Twin Earth assertibility conditions for "T". Both the Earthling and the Twin Earthling will manifest the ability to recognise the presence of what he calls a table, but only we, from our position of Olympian detachment, can say

whether the object in question is a "real" table or merely a "table in the image". But this is no reason to deny that what Earthlings' manifest is the ability to recognise real tables (and not inner representations of tables), unless we adhere to an extreme version of that very Cartesian epistemology criticised by McDowell.

Thus McGinn's contention that externally characterised dispositions to use sentences are not fully manifestable in linguistic practice, and the contention of both McGinn and Devitt that the anti-realist conception of content cannot take account of the environmental embedding of linguistic practice, seem to rest upon the assumption that the anti-realist is committed to something like Hansen's "Lockean epistemology". And there is no more reason to attribute such epistemological views to Dummett than there were found to be, in chapter five above, to attribute such views to Putnam. The anti-realist will readily allow that speakers of a language can recognise that certain external states, of affairs obtain; that they can manifest, in their linguistic practice, an implicit knowledge that the truth conditions for sentences in the language are given by the obtaining of such states of affairs; and that, as a consequence, a theory of meaning for the language can ascribe externally characterised content to sentences of the language. What the anti-realist will not allow, of course, is that speakers can manifest an implicit knowledge of truth conditions whose obtaining is in principle unrecognisable, and that a theory of meaning can ascribe R-Truth conditions to the sentences of language.

Are there, nonetheless, good reasons for thinking that an adequate philosophical account of language should incorporate a realist notion of content? Do we, as McGinn maintains, require the notion of R-Truth

ments for various forms of ontological realism, and (ii) the assumption that the motivation for the programme is primarily epistemological in nature, rather than methodological.

In this work, I have refrained from speculating upon the manner in which an alternative, internalist research programme would construe those features of our linguistic practice that we seek to illuminate in an adequate philosophy of language. More specifically, I have said nothing about the form which should be taken by a non-externalist theory of meaning, nor about the proper formulation for a non-externalist conception of truth. Thus I leave open, in the present context, the question whether we can indeed find Putnam's "narrow path" between metaphysics and relativism - and, indeed, the question whether such a path is the appropriate philosophical route to travel if one departs from the well-worn road of externalism. Such matters - the reader may be thankful or sorrowful! - fall beyond the purview of the present work.

Understanding", the introduction of a realist notion of truth in the metalanguage is not accomplished by the obvious strategy of providing a Tarskian truth-definition, even if we also specify that the meta-language must have a classical logic. Further, as Putnam recognises in his response to Dummett, even "our entire theory" fails to manifest our possession of a realist notion of truth.

VI-d Conclusions

At the beginning of this chapter, I suggested that we might reject
externalism on the grounds that, while not incoherent, it fails to be
helpful in clarifying or resolving philosophically interesting questions
about reference, truth, and language. I then presented arguments to this
effect by Dummett and Wright, and considered three strategies which
defenders of externalism have adopted in the face of these arguments.
I have tried to show that none of these strategies work, and that most
attempts by philosophers to employ such strategies against arguments
for anti-realism rest upon misconceptions of the issues. The conclusion
that I wish to draw is that the onus now rests upon the defenders to
externalism to show that the realist research programme in metaphysics
and the philosophy of language is not bankrupt in the manner suggested.
I am not optimistic that the realist will succeed in this endeavour.
In the meantime, I think the more pressing task is to explore and eval-
uate an alternative philosophical research programme - one that combines
a form of meaning-theoretic anti-realism with an assertibilist notion
of truth.

Let me conclude by briefly recapping the structure of the argument
against externalism developed in the preceding pages. In a sense, the
order of presentation, while satisfying certain dialectical desiderata,
inverts the logical structure of the argument. For we began, in the first
five chapters, with what is essentially the third of the strategies
considered in the present chapter - the attempt to retain the notion of
R-Truth as an indispensable element in a philosophical account of the
functioning of language, while granting the case for meaning-theoretic

anti-realism. We rejected Putnam's claim that the position of one who adopts such a strategy is <u>incoherent</u>, but maintained that his arguments, and the price that the externalist has to pay in order to meet the charge of incoherence, give us good reason to doubt whether such such a strategy is a promising one for philosophers to pursue. For, in the first place, the supposed virtue of externalism as a framework for the philosophy of language - its capacity to explain the "point" of language use, its contribution to the successful prosecution of human projects - is not sacrificed by taking the "internalist turn", but is preserved as part of "scientific" or "empirical" realism - the "sophisticated realism" of the Locke Lectures. Secondly, externalism, but not internalism, divorces reference, meaning, and truth from our practices of <u>ascribing</u> referents, <u>translating</u> between languages and idiolects, and <u>justifying</u> statements. One way of putting this point would be to say that externalism engenders insoluble skeptical problems concerning such practices, and is thus to be rejected on <u>epistemological</u> grounds. But we might put this point another way by saying that the externalist picture fails to answer to our philosophical interests because we are interested in elucidating precisely those practices from which the externalist notions of reference, meaning, and truth are divorced. If we <u>do</u> wish to bring out the epistemological dimensions of externalism, however, we may further note that it obliges us to treat certain "insane" skeptical hypotheses as genuine possibilities - for example, the hypothesis that rocks are conscious. Again, the internalist can simply dismiss such scenarios that traffic in in principle undetectable states of affairs.

Thirdly, metaphysical realism carries with it, not surprisingly,

metaphysical "baggage", such as the idea that Nature itself sorts things into kinds, that Reference is built into Nature itself, and that there is, in Putnam's phrase, a Ready-Made World. We may be prepared to carry such baggage, but it is surely a commitment that should not be undertaken unless it promises to yield some philosophical advantage. Fourthly, our endorsement of meaning-theoretic anti-realism calls into question the very intelligibility of the realist position, for it seems that the grounds for ascribing to us an understanding of the notion of R-Truth in the metalanguage are no better than the grounds for ascribing to us a grasp of recognition-transcendent truth conditions for the sentences in the object language. It might be thought that this does demonstrate the incoherence of the "Reference and Understanding" strategy, for any philosophical position that implies its own unintelligibility is surely incoherent. But we must remember, as I have stressed above, that the arguments for meaning-theoretic anti-realism are methodological in nature, and do not rest upon a proof that we are incapable of grasping R-Truth conditions, or, a fortiori, of grasping the notion of R-Truth.

In the present chapter, I have defended what is taken as a given in both the third realist strategy and in anti-realist criticisms of that strategy, namely, the assumption that one cannot ground semantic realism in meaning-theoretic realism. I have upheld the arguments of Dummett and Wright against representative examples of the strategies adopted by their critics. I have argued that criticisms of Dummett have often rested upon two kinds of mis-readings of the anti-realist programme: (i) a conflation of arguments for semantic realism and argu-

ments for various forms of ontological realism, and (ii) the assumption that the motivation for the programme is primarily epistemological in nature, rather than methodological.

In this work, I have refrained from speculating upon the manner in which an alternative, internalist research programme would construe those features of our linguistic practice that we seek to illuminate in an adequate philosophy of language. More specifically, I have said nothing about the form which should be taken by a non-externalist theory of meaning, nor about the proper formulation for a non-externalist conception of truth. Thus I leave open, in the present context, the question whether we can indeed find Putnam's "narrow path" between metaphysics and relativism - and, indeed, the question whether such a path is the appropriate philosophical route to travel if one departs from the well-worn road of externalism. Such matters - the reader may be thankful or sorrowful! - fall beyond the purview of the present work.

Endnotes - Chapter Six

1. Wright, 1985, pp. 313-314.

2. For an extended treatment of one example of this, see Davies, 1987b.

3. Dummett, 1974, p. 97.

4. Ibid., p. 99.

5. Ibid., p. 101.

6. Loc. cit.

7. Dummett, 1976b, p. 135.

8. Dummett, 1976a, p. 93.

9. Dummett, 1982, p. 55.

10. Loc, cit.

11. Ibid., p. 56.

12. Dummett, 1978, p. xxxviii.

13. Ibid., p. xxxi.

14. Strawson, 1976, p. 17.

15. Dummett, 1978, p. xxv.

16. Ibid., p. xxviii.

17. Ibid., p. xxix.

18. Ibid., p. xxxi.

19. Ibid., p. xxviii.

20. Dummett, 1976a, p. 84.

21. Ibid., p. 87.

22. Ibid., p. 89.

23. Dummett, 1980, p. 11.

24. Ibid., p. 12.

25. Ibid., p. 13.

26. Ibid., p. 16.

27. Dummett, 1969, p. 362.

28. Dummett, 1976a, p. 77.

29. Ibid., p. 78.

30. Ibid., p. 99.

31. Ibid., pp. 100-101.

32. Ibid., p. 101

33. Dummett, 1976b, pp. 134-135.

34. Dummett, 1976a, p. 80.

35. Loc. cit.

36. Ibid., p. 81.

37. Loc. cit.

38. Loc. cit.

39. Ibid., p. 82.

40. Prawitz, 1980.

41. See especially Dummett, 1976b, pp. 134-135.

42. Prawitz, 1980, p. 5

43. Prawitz, 1980, p. 4.

44. Prawitz, 1980, p. 5.

45. Loc. cit.

46. Wright, 1980, p. 125.

47. Loc. cit.

48. Dummett, 1980, p. 15.

49. Ibid., p. 16.

50. Devitt, 1983b.

51. Ibid., p. 74.

52. Ibid., pp. 74-75.

53. Ibid., p. 76.

54. Ibid., pp. 76-77.

55. Ibid., p. 83.

56. Ibid., p. 84.

57. Ibid., p. 85.

58. Ibid., p. 84.

59. Ibid., p. 83.

60. Ibid., p. 86.

61. Ibid., p. 90.

62. Ibid., p. 91.

63. Ibid., p. 92.

64. A similar point is made in George, 1984.

65. McGinn, 1982a

66. Devitt, 1983b, p. 84.

67. Ibid., p. 92.,

68. Dummett, 1976a, p. 70...

69. Devitt, 1983b, p. 85.

70. Dummett, 1976a, pp. 69-70.

71. Wright, 1985, p. 313.

72. Devitt, 1983b, p. 98.

73. Ibid., p. 97.

74. See Devitt, 1983b, section II-3.

75. Fodor, 1968, chapter 4; 1981, Introduction.

76. Devitt, 1983b, pp. 92-93.

77. George, 1984.

78. See Putnam, 1976b.

79. Horwich, 1982.

80. See, for example, Dummett, 1978, p. xxxviii.

81. Wright, 1976, p. 227.

82. McGinn, 1980.

83. Ibid., p. 20.

84. Ibid., p. 21.

85. Ibid., p. 21; p. 24.

86. Ibid., p. 25.

87. Demopoulos, 1982, p. 148.

88. McGinn, 1980, p. 28.

89. Ibid., p. 30.

90. Dummett, 1976a, pp. 110-111.

91. Demopoulos, 1982, pp. 148-149.

92. See also George, 1984, on this point.

93. Wright, 1976, p. 245.

94. Strawson, 1976, p. 17.

95. Wright, 1979, p. 286.

96. Strawson, 1976.

97. Ibid., p. 18.

98. Loc. cit.

99. Ibid., p. 19.

100. Dummett, 1978, p. xxxvi.

101. Wright, 1979, pp. 287-288.

102. Dummett, 1978, p. xxxv.

103. Wright, 1979, p. 289.

104. Strawson, 1976, p. 16.

105. Wright, 1979, p. 294.

106. Ibid., p. 289.

107. Loc. cit.

108. Ibid., p. 290.

109. Dummett, 1973b, p. 216, cited in Blackburn, 1984, p. 65.

110. Blackburn, 1984, p.)65.

111. Ibid., p. 66.

112. Wright, 1985, p. 13.

113. Loc, cot.

114. McDowell, 1981, p. 227.

115. Ibid., p. 228.

116. Ibid., p. 240.

117. Ibid., pp. 228-229.

118. Ibid., p. 229.

119. Ibid., p. 230.

120. Ibid., p. 225.

121. Ibid., p. 231.

122. See Horwich, 1982, for a similar strategy. See Davies, 1987b, for
     a critical discussion of Horwich's paper. The claim that we manifest
     a realist understanding of sentences by using those sentences to
     make realist assertions is similar to McGinn's contention (McGinn,
     1980), cited earlier, that speakers manifest a knowledge of R-Truth
     conditions by interpreting other's assertions as having realist
     content. McDowell distances himself from McGinn's strategy, however:
     see McDowell, 1981, p. 240. fn. 26.

123. McDowell, 1981, p. 230.

124. Ibid., pp. 233-236.

125. See McDowell, 1981, p. 234, fn. 18, for an acknowledgement of this
     point.

126. McDowell, 1981, p. 237.

127. Loc. cit.

128. Ibid., p. 238.

129. Ibid., p. 248.

130. Ibid., p. 239.

131. Ibid., p. 225.

132. Ibid., p. 239.

133. Loc. cit.

134. Ibid., pp. 242-243.

135. Ibid., p. 243.

136. Ibid., p. 244.

137. Dummett, 1974.

138. Ibid., p. 97.

139. Ibid., p. 99.

140. Dummett, 1973b, p. 219.

141. McDowell, 1981, p. 225.

142. Ibid., p. 239.

143. Loc. cit.

144. Dummett, 1976a, p. 70; 1973a, pp. 679-681.

145. Dummett, 1976a, p. 70.

146. McDowell, 1981, p. 248.

147. Loc. cit.

148. Ibid., p) 238.

149. McGinn, 1982b.

150. Ibid., p. 114-115.

151. Devitt, 1983b, p. 91.

152. McGinn, 1982b, p. 129.

153. Ibid., p. 130.

154. Devitt, 1983b, p. 95.

155. McGinn, 1982b, p. 130.

156. Ibid., p, 116.

157. Ibid., p. 131.

158. Ibid., p. 118.

159. Ibid., p. 119.

160. Ibid., p. 125.

161. Ibid., p. 127.

162. Putnam, 1975c.

163. McGinn, 1982b, p. 131.

Bibliography

Blackburn, Simon, 1984. <u>Spreading The Word</u>. Oxford: Clarendon, 1984.

Brueckner, Anthony, 1984. "Putnam's Model-Theoretic Argument Against
   Metaphysical Realism." <u>Analysis</u>, 44 (1984), pp. 134-140.

----------. "Brains In A Vat." <u>Journal of Philosophy</u>, LXXXIII (1986),
   pp. 148-167.

Currie, Gregory, 1982. "A Note On Realism." <u>Philosophy of Science</u>,
   49 (1982), pp. 263-267.

Davidson, Donald, 1967. "Truth and Meaning." <u>Synthese</u>, XVII. (1967),
   pp. 309-322.

Davies, David, 1987a. "How Not to OutSmart the Anti-Realist."
   <u>Analysis</u>, 47 (1987), pp. 1-8.

----------., 1987b. "Horwich on Semantic and Metaphysical Realism."
   <u>Philosophy of Science</u>, 54 (1987), forthcoming.

Demopoulos, William, 1982. "The Rejection of Truth-Conditional Sem-
   antics by Dummett and Putnam." <u>Philosophical Topics</u>, 13, No. 1
   (1982), pp. 135-153.

----------, and Michael Friedman, 1985. "Critical Notice: Bertrand
   Russell's <u>The Analysis of Matter</u>: Its Historical Context and Con-
   temporary Interest." <u>Philosophy of Science</u>, 52 (1985), pp. 621-
   639.

Devitt, Michael, 1983a. "Realism and the Renegade Putnam." <u>Nous</u>,
   XVII (1983), pp. 291-301.

----------., 1983b. "Dummett's Anti-Realism." <u>Journal of Philosophy</u>,
   LXXX (1983), pp. 73-99.

Dummett, Michael, 1969. "The Reality of the Past". Reprinted in
Dummett, 1978, pp. 358-374.

----------., 1973a. Frege: Philosophy of Language. London: Duckworth,
1973.

----------., 1973b. "The Philosophical Basis of Intuitionist Logic."
Reprinted in Dummett, 1978, pp. 215-247.

----------., 1974. "What is a Theory of Meaning? I." In Guttenplan,
1975, pp. 97-138.

----------., 1976a. "What is a Theory of Meaning? II." In Evans and
McDowell, 1976, pp. 67-137.

----------., 1976b. "What Does the Appeal to Use Do For the Theory
of Meaning?" In Margalit, 1979, pp. 123-135.

----------., 1976c. "Comments on Putnam's 'Reference and Understand-
ing'." In Margalit, 1979, pp. 218-225.

----------., 1978. Truth and Other Enigmas. Cambridge: Harvard Univ-
ersity Press, 1978.

----------., 1980. "Comments on Professor Prawitz's Paper." In von
Wright, 1980, pp. 11-18.

----------., 1982. "Realism." Synthese, 52 (1982), pp. 55-112.

Evans, Gareth, and John McDowell, eds., 1976. Truth and Meaning.
Oxford: Clarendon Press, 1976.

Feldman, Susan, 1984. "Refutation of Dogmatism: Putnam's Brains in
Vats." Southern Journal of Philosophy, 22 (1984), pp. 323-329.

Field, Hartry, 1972. "Tarski's Theory of Truth." Journal of Phil-
osophy, LXIX (1972), pp. 347-375.

----------., 1982. "Realism and Relativism." Journal of Philosophy,

LXXIX (1982), pp. 553-567.

Fodor, Jerry, 1968. Psychological Explanation. New York: Random House, 1968.

----------., 1981. Representations. Cambridge: M.I.T. Press, 1981.

George, Alexander, 1984. "On Devitt on Dummett." Journal of Philosophy, LXXXI (1984), pp.

Glymour, Clark, 1982. "Conceptual Scheming." Synthese, 51 (1982), pp. 169-180.

Goldman, Alan H., 1984. "Fanciful Arguments for Realism." Mind, XCIII (1984), pp. 19-38.

Goldman, Alvin, 1986. Epistemology and Cognition. Cambridge: Harvard University Press, 1986.

Goodman, Nelson, 1954. Fact, Fiction, and Forecast. Indianapolis: Bobbs-Merrill, 1954.

----------., 1960. "The Way The World Is." Reprinted in Goodman, 1972, pp. 24-32.

----------., 1968. Languages of Art. Indianapolis: Bobbs-Merrill, 1968.

----------., 1972. Problems and Projects. Indianapolis: Hackett., 1972.

----------., 1978. Ways of Worldmaking. Indianapolis: Hackett, 1978.

Guttenplan, Samuel, ed., 1975. Mind and Language. Oxford: Clarendon Press, 1975.

Hacking, Ian, 1983. Representing and Intervening. Cambridge: Cambridge University Press, 1983.

Hansen, Carsten, 1987. "Putnam's Indeterminacy Argument: The Skol-

emisation of Absolutely Everything." Philosophical Studies,

  51 (1987), pp. 77-99.

Harrison, Jonathan, 1985. "Professor Putnam on Brains in Vats."

  Erkenntnis, (1985), pp. 55-57.

Holtzman, Stephen, and Christopher Leich, eds., 1981. Wittgenstein:

  To Follow a Rule. London: Routledge and Kegan Paul, 1981.

Horwich, Paul, 1982. "Three Forms of Realism." Synthese, 51 (1982),

  pp. 181-202.

Kaplan, David, 1977. "Demonstratives - draft #2." Unpublished.

Koethe, John, 1979. "Putnam's Argument Against Realism." Philo-

  sophical Review, LXXXVIII (1979), pp. 92-99.

Kripke, Saul, 1972. "Naming and Necessity." In Donald Davidson and

  Gilbert Harman, eds., The Semantics of Natural Language,

  (Dordrecht: Reidel, 1972).

Lewis, David, 1984. "Putnam's Paradox." Australasian Journal of

  Philosophy, 62 (1984), pp. 221-236.

Margalit, A., ed., 1979. Meaning and Use. Dordrecht: Reidel, 1979.

McDowell, John, 1981. "Anti-Realism and the Epistemology of Under-

  standing." In Herman Parrett and Jacques Bouveresse, eds., Mean-

  ing and Understanding, (de Guyter, 1981), pp. 225-248.

McGinn, Colin, 1980. "Truth and Use." In Platts, 1980, pp. 19-40.

------., 1982a. "Two Notions of Realism?" Philosophical Topics,

  13, No. 1 (1982), pp. 123-134.

----------., 1982b. "Realist Semantics and Content Ascription."

  Synthese, 52 (1982), pp. 113-134.

McIntyre, Jane, 1984. "Putnam's Brains." Analysis, 44 (1984), pp.

  59-61.

Melchert, Norman, 1986. "Metaphysical Realism and History." Analysis,

    46 (1986), pp. 36-38.

Merrill, G.H., 1980. "The Model-Theoretic Argument Against Realism."

    Philosophy of Science, 47 (1980), pp. 69-81.

Nagel, Thomas, 1974. "What Is It Like to Be a Bat?" Philosophical

    Review, 83 (1974), pp. 435-450.

Pearce, D., and V. Rantala, 1982. "Realism and Reference: Some Comm-

    ents on Putnam." Synthese, 52 (1982), pp. 39-53.

Plantinga, Alvin, 1982. "How to Be an Anti-Realist." Proceedings

    and Addresses of the American Philosophical Association, 56

    (1982), pp. 47-70.

Platts, Mark, 1980. Reference, Truth, and Reality. London: Routledge,

    Kegan Paul, 1980.

Prawitz, Dag, 1980. "Intuitionist Logic: A Philosophical Challenge."

    In von Wright, 1980, pp. 1-10.

Putnam, Hilary, 1960. "Do True Assertions Correspond to Reality?"

    In Putnam, 1975b, pp. 70-84.

----------., 1962. "The Analytic and the Synthetic." In Putnam,

    1975b, pp. 33-69.

----------., 1971. "Philosophy of Logic." In Putnam 1975a, pp..

    323-357.

----------., 1975a. Philosophical Papers, Volume One. Cambridge:

    Cambridge University Press, 1975.

----------., 1975b. Philosophical Papers, Volume Two. Cambridge:

    Cambridge University Press, 1975.

----------., 1975c. "The Meaning of 'Meaning'." In Putnam, 1975b,

pp. 215-271.

----------., 1975d. "Literature, Science, and Reflection." Reprinted in Putnam, 1978, pp. 83-94.

----------., 1975e. "What Is Realism?" Proceedings of the Aristotelean Society, 76 (1975-76), pp. 177-194.

----------., 1976a. The Locke Lectures: Meaning and Knowledge. In Putnam, 1978, pp. 7-80.

----------., 1976b. "Reference and Understanding." In Putnam, 1978, pp. 97-117.

----------., 1976c. "Reply to Dummett." In Margalit, 1979, pp. 226-228.

----------., 1976d. "Realism and Reason." In Putnam, 1978, pp. 123-140.

----------., 1978. Meaning and the Moral Sciences. London: Routledge and Kegan Paul, 1978.

----------., 1979. "Reflections on Goodman's Ways of Worldmaking." Reprinted in Putnam, 1983a, pp. 155-169.

----------., 1980. "Models and Reality." Journal of Symbolic Logic, XLV (1980), pp. 464-482.

----------., 1981a. Reason, Truth, and History. Cambridge: Cambridge University Press, 1981.

----------., 1981b. "Quantum Mechanics and the Observer." Reprinted in Putnam, 1983a, pp. 248-270.

----------., 1982a. "Reply to Two Realists." Journal of Philosophy, LXXIX (1982), pp. 575-577.

----------., 1982b. "Why There Isn't a Ready-Made World." In Putnam, 1983a, pp. 205-228.

----------., 1982c. "Why Reason Can't Be Naturalised." In Putnam,
   1983a, pp. 229-247.

----------., 1982d. "Three Kinds of Scientific Realism." Philo-
   sophical Quarterly, 32 (1982), pp. 195-200.

----------., 1983a. Philosophical Papers: Volume Three. Cambridge:
   Cambridge University Press, 1983.

----------., 1983b. "Vagueness and Alternative Logic." In Putnam,
   1983a, pp. 271-286.

----------., 1983c. "Beyond Historicism." In Putnam, 1983a, pp.
   287-303.

----------., 1983d. "Reference and Truth." In Putnam, 1983a, pp.
   69-86.

----------., 1983e. "On Truth." In Leigh Cauman et. al., eds.,
   How Many Questions?, (Indianapolis: Hackett, 1983), pp. 35-56.

----------., 1985a. "Why Is a Philosopher?" Unpublished.

----------., 1985b. The Carus Lectures: The Many Faces of Realism.
   Forthcoming.

Quine, Willard van Orman, 1960. Word and Object. Cambridge: M.I.T.
   Press, 1960.

----------., 1969. "Ontological Relativity." In Ontological Rel-
   ativity and Other Essays, (New York: Columbia University Press,
   1969), pp. 26-68.

Rorty, Richard, 1972. "The World Well Lost." Journal of Philosophy,
   LXIX (1972), pp. 649-665.

----------., 1979. Philosophy and the Mirror of Nature. Princeton:
   Princeton University Press, 1979.

Smart, J.J.C., 1982. "Metaphysical Realism." Analysis, 42 (1982), pp. 1-3.

——————., 1986. "Realism v. Idealism." Philosophy, 61 (1986), pp. 295-312.

Smith, Peter, 1983. "Smart's Argument for Realism." Analysis, 43 (1983), pp. 74-78.

——————., 1984. "Could We Be Brains in a Vat?" Canadian Journal of Philosophy, XIV (1984), pp. 115-123.

——————., 1986. "Metaphysical Realism and Historical Interpretation." Analysis, 46 (1986), pp. 157-158.

Stephens, James, and Lily-Marlene Russow, 1985. "Brains in Vats and the Internalist Perspective." Australasian Journal of Philosophy, 63 (1985), pp. 205-212.

Strawson, Peter, 1976. "Scruton and Wright on Anti-Realism, Etc." Proceedings of the Aristotelean Society (New Series), 77 (1976-1977), pp. 15-22.

Tarski, Alfred, 1951. "The Concept of Truth in Formalised Languages." Reprinted in his Logic, Semantics, Metamathematics (Oxford: Oxford University Press, 1956), pp. 152-278.

von Wright, G.H., 1980. Logic and Philosophy. The Hague: Martinus Nijhoff, 1980.

Wright, Crispin, 1976. "Truth-Conditions and Criteria." Proceedings of the Aristotelean Society, Supplementary Volume L (1976), pp. 196-245.

——————., 1979. "Strawson on Anti-Realism." Synthese, 40 (1979), pp. 283-299.

——————., 1980. Wittgenstein on the Foundations of Mathematics.

London: Duckworth, 1980.

----------., 1981a   "Rule-Following: Objectivity and the Theory of
Meaning." In Holtzman and Leich, 1981, pp. 99-117.

----------., 1981b. "Dummett and Revisionism." Philosophical Quar-
terly, 31 (1981), pp. 47-67.

----------., 1985. "Review of Simon Blackburn, Spreading the Word."
Mind, XCIV (1985), pp. 311-318.