THE JOURNAL OF SYMBOLIC LOGIC
Volume 45, Number 3, Sept. 1980

MODELS AND REALITY[1]

HILARY PUTNAM

In 1922 Skolem delivered an address before the Fifth Congress of Scandinavian Mathematicians in which he pointed out what he called a "relativity of set-theoretic notions". This "relativity" has frequently been regarded as paradoxical; but today, although one hears the expression "the Löwenheim-Skolem Paradox", it seems to be thought of as only an *apparent* paradox, something the cognoscenti enjoy but are not seriously troubled by. Thus van Heijenoort writes, "The existence of such a 'relativity' is sometimes referred to as the Löwenheim-Skolem Paradox. But, of course, it is not a paradox in the sense of an antinomy; it is a novel and unexpected feature of formal systems." In this address I want to take up Skolem's arguments, not with the aim of refuting them but with the aim of extending them in somewhat the direction he seemed to be indicating. It is not my claim that the "Löwenheim-Skolem Paradox" is an antinomy *in formal logic*; but I shall argue that it *is* an antinomy, or something close to it, in *philosophy of language*. Moreover, I shall argue that the resolution of the antinomy—the only resolution that I myself can see as making sense—has profound implications for the great metaphysical dispute about realism which has always been the central dispute in the philosophy of language.

The structure of my argument will be as follows: I shall point out that in many different areas there are three main positions on reference and truth: there is the extreme Platonist position, which posits nonnatural mental powers of directly "grasping" forms (it is characteristic of this position that "understanding" or "grasping" is itself an irreducible and unexplicated notion); there is the verificationist position which replaces the classical notion of truth with the notion of verification or proof, at least when it comes to describing how the language is understood; and there is the moderate realist position which seeks to preserve the centrality of the classical notions of truth and reference without postulating nonnatural mental powers. I shall argue that it is, unfortunately, the *moderate* realist position which is put into deep trouble by the Löwenheim-Skolem Theorem and related model-theoretic results. Finally I will opt for verificationism as a way of preserving the outlook of scientific or empirical realism, which is totally jettisoned by Platonism, even though this means giving up *metaphysical* realism.

The Löwenheim-Skolem Theorem says that a satisfiable first-order theory (in

Received January 30, 1978

[1]Presidential Address delivered before the Winter Meeting of the Association for Symbolic Logic in Washington, D. C., December 29, 1977. I wish to thank Bas van Fraassen for valuable comments on and criticisms of an earlier version.

0022-4812/80/4503-0004/$05.75

a countable language) has a countable model. Consider the sentence:

(i) $-(ER)(R$ is *one-to-one. The domain of* $R \subset N$. *The range of values of R is* $S)$ where 'N' is a formal term for the set of all whole numbers and the three conjuncts in the matrix have the obvious first-order definitions.

Replace 'S' with the formal term for the set of all real numbers in your favorite formalized set theory. Then (i) will be a *theorem* (proved by Cantor's celebrated "diagonal argument"). So your formalized set theory *says* that a certain set (call it "S") is nondenumerable. So S must *be* nondenumerable in all *models* of your set theory. So your set theory—say ZF (Zermelo-Fraenkel set theory) has only nondenumerable models. But this is impossible! For, by the Löwenheim-Skolem Theorem, *no* theory can have *only* nondenumerable models; if a theory has a nondenumerable model, it must have denumerably infinite ones as well. Contradiction.

The resolution of this apparent contradiction is not hard, as Skolem points out (and it is not this apparent contradiction that I referred to as an antinomy, or close to an antinomy). For (i) only "says that S is nondenumerable when the quantifier (ER) is interpreted as ranging over *all* relations on $N \times S$. But when we pick a *denumerable* model for the language of set theory, "(ER)" does not range over *all* relations; it ranges only over relations *in the model.* (i) only "says" that S is nondenumerable in a *relative* sense: the sense that the members of S cannot be put in one-to-one correspondence with a subset of N by any R *in the model.* A set S can be "nondenumerable" in this *relative* sense and yet be denumerable "in reality". This happens when there *are* one-to-one correspondences between S and N but all of them lie outside the given model. What is a "countable" set from the point of view of one model may be an uncountable set from the point of view of another model. As Skolem sums it up, "even the notions 'finite', 'infinite', 'simply infinite sequence' and so forth turn out to be merely relative within axiomatic set theory".

The philosophical problem. Up to a point all commentators agree on the significance of the existence of "unintended" interpretations, e.g., models in which what are "supposed to be" nondenumerable sets are "in reality" denumerable. All commentators agree that the existence of such models shows that the "intended" interpretation, or, as some prefer to speak, the "intuitive notion of a set", is not "captured" by the formal system. But if *axioms* cannot capture the "intuitive notion of a set", what possibly could?

A technical fact is of relevance here. The Löwenheim-Skolem Theorem has a strong form (the so-called "downward Löwenheim-Skolem Theorem"), which requires the axiom of choice to prove, and which tells us that a satisfiable first-order theory (in a countable language) has a countable model which is a submodel of any given model. In other words if we are given a nondenumerable model M for a theory, then we can find a countable model M' of that same theory in which the predicate symbols stand for the same relations (restricted to the smaller universe in the obvious way) as they did in the original model. The only difference between M and M' is that the "universe" of M'—i.e., the totality that the variables of quantification range over— is a proper subset of the "universe" of M.

Now the argument that Skolem gave, and that shows that "the intuitive notion of a set" (if there is such a thing) is not "captured" by any formal system, shows that even a *formalization of total science* (if one could construct such a thing), or even a *formalization of all our beliefs* (whether they count as "science" or not), could not rule out denumerable interpretations, and, *a fortiori*, such a formalization could not rule out *unintended* interpretations of this notion.

This shows that "theoretical constraints", whether they come from set theory itself or from "total science", cannot fix the interpretation of the notion *set* in the "intended" way. What of "operational constraints"?

Even if we allow that there might be a *denumerable infinity* of measurable "magnitudes", and that each of them might be measured to *arbitrary rational accuracy* (which certainly seems a utopian assumption), it would not help. For, by the "downward Löwenheim-Skolem Theorem", we can find a countable submodel of the "standard" model (if there is such a thing) in which countably many predicates (each of which may have countably many things in its extension) have their extensions preserved. In particular, we can fix the values of countable many magnitudes at all rational space-time points, and still find a countable submodel which meets all the constraints. In short, there certainly seems to be a *countable* model of our *entire body of belief* which meets all operational constraints.

The philosophical problem appears at just this point. If we are told, "axiomatic set theory does not capture the intuitive notion of a set", then it is natural to think that *something else*—our "understanding"—does capture it. But what can our "understanding" come to, at least for a naturalistically minded philosopher, which is more than *the way we use our language*? The Skolem argument can be extended, as we have just seen, to show that the *total use of the language* (operational plus theoretical constraints) does not "fix" a unique "intended interpretation" any more than axiomatic set theory by itself does.

This observation can push a philosopher of mathematics in two different ways. If he is inclined to Platonism, he will take this as evidence that the mind has mysterious faculties of "grasping concepts" (or "perceiving mathematical objects") which the naturalistically minded philosopher will never succeed in giving an account of. But if he is inclined to some species of verificationism (i.e., to indentifying truth with verifiability, rather than with some classical "correspondence with reality") he will say, "Nonsense! All the 'paradox' shows is that our understanding of 'The real numbers are nondenumerable' consists in our knowing *what it is for this to be proved*, and not in our 'grasp' of a 'model'." In short, the extreme positions—Platonism and verificationism—seem to receive comfort from the Löwenheim-Skolem Parodox; it is only the "moderate" position (which tries to avoid mysterious "perceptions" of "mathematical objects" while retaining a classical notion of truth) which is in deep trouble.

An epistemological/logical digression. The problem just pointed out is a serious problem for any philosopher or philosophically minded logician who wishes to view set theory as the description of a determinate independently existing reality. But from a mathematical point of view, it may appear immaterial: what does it matter if there are many different models of set theory, and not a unique "intended model" *if they all satisfy the same sentences*? What we want to know as mathematicians is what sentences of set theory are true; we do not want to have the sets themselves in our hands.

Unfortunately, the argument can be extended. First of all, the theoretical constraints we have been speaking of must, on a naturalistic view, come from only two sources: they must come from something like human decision or convention, whatever the source of the "naturalness" of the decisions or conventions may be, or from human experience, both experience with nature (which is undoubtedly the source of our most basic "mathematical intuitions", even if it be unfashionable to say so), and experience with "doing mathematics". It is hard to believe that either or both of these sources together can ever give us a *complete* set of axioms for set theory (since, for one thing, a complete set of axioms would have to be nonrecursive, and it is hard to envisage coming to have a nonrecursive set of axioms in the literature or in our heads even in the unlikely event that the human race went on forever doing set theory); and if a complete set of axioms is impossible, and the intended model*s* (in the plural) are singled out only by theoretical plus operational constraints then sentences which are independent of the axioms which we will arrive at in the limit of set-theoretic inquiry really have *no* determinate truth value; they are just true in some intended models and false in others.

To show what bearing this fact may have on actual set-theoretic inquiry, I will have to digress for a moment into technical logic. In 1938 Gödel put forward a new axiom for set theory: the axiom "$V = L$". Here L is the class of all constructible sets, that is, the class of all sets which can be defined by a certain constructive procedure if we pretend to have names available for all the ordinals, however large. (Of course, this sense of "constructible" would be anathema to constructive mathematicians.) V is the universe of all sets. So "$V = L$" just says *all sets are constructible*. By considering the inner model for set theory in which "$V = L$" is true, Gödel was able to prove the relative consistency of ZF and ZF *plus* the axiom of choice and the generalized continuum hypothesis.

"$V = L$" is certainly an important sentence, mathematically speaking. Is it *true*?

Gödel briefly considered proposing that we *add* "$V = L$" to the accepted axioms for set theory, as a sort of meaning stipulation, but he soon changed his mind. His later view was that "$V = L$" is *really* false, even though it is consistent with set theory, if set theory is itself consistent.

Gödel's intuition is widely shared among working set theorists. But does this "intuition" make sense?

Let *MAG* be a countable set of physical magnitudes which includes all magnitudes that sentient beings in this physical universe can actually measure (it certainly seems plausible that we cannot hope to measure more than a countable number of physical magnitudes). Let *OP* be the "correct" assignment of values; that is, the assignment which assigns to each member of *MAG* the value that that magnitude actually has at each rational space-time point. Then all the information "operational constraints" might give us (and, in fact, infinitely more) is coded into *OP*.

One technical term: an *ω-model* for a set theory is a model in which the *natural numbers* are ordered as they are "supposed to be"; that is, the sequence of "natural numbers" of the model is an ω-sequence.

Now for a small theorem.[2]

THEOREM. *ZF plus $V = L$ has an ω-model which contains any given countable set of real numbers.*

PROOF. Since a countable set of reals can be coded as a single real by well-known techniques, it suffices to prove that *for every real s, there is an M such that M is an ω-model for ZF plus $V = L$ and s is represented in M.*

By the "downward Löwenheim-Skolem Theorem", this statement is true if and only if the following statement is:

For every real s, there is a countable M such that M is an ω-model for ZF plus $V = L$ and s is represented in M.

Countable structures with the property that the "natural numbers" of the structure form an ω-sequence can be coded as reals by standard techniques. When this is properly done, the predicate "*M is an ω-model for ZF plus $V = L$ and s is represented in M*" becomes a two-place *arithmetical* predicate of *reals M, s.* The above sentence thus has the logical form (*for every real s*) (*there is a real M*) $(\cdots M, s, \cdots)$. In short, the sentence is a Π_2-sentence.

Now, consider this sentence *in the inner model $V = L$.* For every *s in the inner model*—that is, for every *s* in *L*—there is a model—namely *L* itself—which satisfies "$V = L$" and contains *s*. By the downward Löwenheim-Skolem Theorem, there is a countable submodel which is elementary equivalent to *L* and contains *s*. (Strictly speaking, we need here not just the downward Löwenheim-Skolem Theorem, but the "Skolem hull" construction which is used to prove that theorem.) By Gödel's work, this countable submodel itself lies in *L*, and as is easily verified, so does the real that codes it. So the above Π_2-sentence is true in the inner model $V = L$.

But Schoenfield has proved that Π_2-sentences are *absolute:* if a Π_2-sentence is true in *L*, then it must be true in *V*. So the above sentence is true in *V*. □

What makes this theorem startling is the following reflection: suppose that Gödel is right, and "$V = L$" is *false* ("in reality"). Suppose that there is, in fact, a *nonconstructible real number* (as Gödel also believes). Since the predicate "is constructible" is absolute in *β-models*—that is, in models in which the "wellorderings" *relative to the model* are wellorderings "in reality" (recall Skolem's "relativity of set-theoretic notions"!), no model containing such a nonconstructible *s* can satisfy "*s* is constructible" and be a *β-model.* But, by the above theorem, a model containing *s can* satisfy "*s* is constructible" (because it satisfies "$V = L$", and "$V = L$" says *everything* is constructible) and be an *ω-model.*

Now, suppose we formalize *the entire language of science* within the set theory ZF *plus $V = L$*. Any model for ZF which contains an abstract set isomorphic to *OP* can be extended to a model for this formalized language of science which is *standard with respect to OP*—hence, even if *OP* is nonconstructible "in reality", we can find a model *for the entire language of science* which satisfies *everything is constructible* and which assigns the correct values to all the physical magnitudes in *MAG* at all rational space-time points.

[2] Barwise has proved the much stronger theorem that every countable model of ZF has a proper end extension which is a model of ZF + $V = L$ (in *Infinitary methods in the model theory of set theory*, published in ***Logic Colloquium* '69**). The theorem in the text was proved by me before 1963.

The claim Gödel makes is that "$V = L$" is false "in reality". But what on earth can this mean? It must mean, at the very least, that in the case just envisaged, the model we have described in which "$V = L$" holds would not be *the intended model.* But why not? It satisfies all theoretical constraints; and we have gone to great length to make sure it satisfies all operational constraints as well.

Perhaps someone will say that "$V \neq L$" (or something which implies that V does not equal L) should be added to the axioms of ZF as an additional "theoretical constraint". (Gödel often speaks of new axioms someday becoming evident.) But, while this may be acceptable from a nonrealist standpoint, it can hardly be acceptable from a realist standpoint. For the realist standpoint is that there is *a fact of the matter*—a fact independent of our legislation—as to whether $V = L$ or not. A realist like Gödel holds that we have access to an "intended interpretation" of ZF, where the access is not simply by linguistic stipulation.

What the above argument shows is that if the "intended interpretation" is fixed only by theoretical plus operational constraints, then if "$V \neq L$" does not follow from those theoretical constraints—if we do not *decide* to make $V = L$ true or to make $V = L$ false—then there will be "intended" models in which $V = L$ is *true.* If I am right, then the "relativity of set-theoretic notions" extends to a *relativity of the truth value of* "$V = L$" (and, by similar arguments, of the axiom of choice and the continuum hypothesis as well).

Operational constraints and counterfactuals. It may seem to some that there is a major equivocation in the notion of what *can* be measured, or observed, which endangers the apparently crucial claim that the evidence we *could* have amounts to at most denumerably many facts. Imagine a measuring apparatus that simply detects the presence of a particle within a finite volume dv around its own geometric center during each full minute on its clock. Certainly it comes up with at most denumerably many reports (each *yes* or *no*) even if it is left to run forever. But how many are the facts it *could* report? Well, if it were jiggled a little, by chance let us say, its geometric center would shift r centimeters in a given direction. It would then report totally different facts. Since for each number r it could be jiggled that way, the number of reports it could produce is nondenumerable—and it does not matter to this that we, and the apparatus itself, are incapable of distinguishing every real number r from every other one. The problem is simply one of scope for the modal word "can". In my argument, I must be identifying what I call operational constraints, not with the totality of facts that could be registered by observation—i.e., ones that either will be registered, or would be registered if certain chance perturbations occurred—but with the totality of facts that will in actuality be registered or observed, whatever those be.

In reply, I would point out that even if the measuring apparatus *were* jiggled r centimeters in a given direction, we could only know the real number r to some rational approximation. Now, if the intervals involved are all rational, there are only *countably* many facts of the form: *if action A* (an action described with respect to place, time, and character up to some finite "tolerance") were performed, then the result $r \pm \varepsilon$ (a result described up to some rational tolerance) *would be obtained with probability in the interval a, b.* To know all facts of this form would be to know the *probability distribution* of all possible observable results of all possible actions. Our argument shows that a model could be constructed which agrees with all of these facts.

There is a deeper point to be made about this objection, however. Suppose we "first orderize" counterfactual talk, say, by including *events* in the ontology of our theory and introducing a predicate ("subjunctively necessitates") for the counterfactual connection between unactualized event types at a given place-time. Then our argument shows that a model exists which fits all the facts that will actually be registered or observed and fits our theoretical constraints, and this model *induces* an interpretation of the counterfactual idiom (a "similarity metric on possible worlds", in David Lewis' theory) which renders true just the counterfactuals that are true according to some completion of our theory. Thus appeal to counterfactual observations cannot rule out any models at all unless the interpretation of the counterfactual idiom itself is *already* fixed by something beyond operational and theoretical constraints.

(A related point is made by Wittgenstein in his *Philosophical Investigations*: talk about what an ideal machine—or God—could compute is talk *within* mathematics—in disguise—and cannot serve to fix the interpretation of mathematics. "God", too, has many interpretations.)

"Decision" and "convention". I have used the word "decision" in connection with open questions in set theory, and obviously this is a poor word. One cannot simply sit down in one's study and "decide" that "$V = L$" is to be true, or that the axiom of choice is to be true. Nor would it be appropriate for the mathematical community to call an international convention and legislate these matters. Yet, it seems to me that if we encountered an extra-terrestrial species of intelligent beings who had developed a high level of mathematics, and it turned out that they *rejected* the axiom of choice (perhaps because of the Tarski-Banach Theorem[3]), it would be wrong to regard them as simply making a *mistake*. To do *that* would, on my view, amount to saying that acceptance of the axiom of choice is built into our notion of rationality itself; that does not seem to me to be the case. To be sure, our acceptance of choice is not arbitrary; all kinds of "intuitions" (based, most likely, on experience with the finite) support it; its mathematical fertility supports it; but none of this is *so* strong that we could say that an equally successful culture which based *its* mathematics on principles *incompatible* with choice (e.g., on the so-called "axiom of determinacy"[4]) was *irrational*.

[3] This is a very counterintuitive consequence of the axiom of choice. Call two objects A, B "congruent by finite decomposition" if they can be divided into finitely many disjoint point sets $A_1, \ldots, A_n, B_1, \ldots B_n$, such that $A = A_1 \cup A_2 \cup \cdots \cup A_n$, $B = B_1 \cup B_2 \cup \cdots \cup B_n$, and (for $i = 1, 2, \ldots, n$) A_i is congruent to B_i. Then Tarski and Banach showed that *all spheres are congruent by finite decomposition.*

[4] This axiom, first studied by J. Mycielski (*On the axiom of determinacy*", ***Fundamenta Mathematicae***, 1963) asserts that infinite games with perfect information are determined, i.e. there is a winning strategy for either the first or second player. AD (the axiom of determinacy) implies the existence of a nontrivial countably additive two-valued measure on the real numbers, contradicting a well-known consequence of the axiom of choice.

But if both systems of set theory—ours and the extra-terrestrials'—count as *rational*, what sense does it make to call one *true* and the others *false*? From the Platonist's point of view there is no trouble in answering this question. "The axiom of choice is true—true in *the* model", he will say (if he believes the axiom of choice). "We are right and the extra-terrestrials are wrong." But what is *the* model? If the intended model is singled out by theoretical and operational constraints, then, first, "the" intended model is plural not singular (so the "the" is inappropriate—our theoretical and operational constraints fit many models, not just one, and so do those of the extra-terrestrials as we saw before. Secondly, the intended models for us do satisfy the axiom of choice and the extra-terrestrially intended models do not; we are not talking about the same models, so there is no question of a "mistake" on one side or the other.

The Platonist will reply that what this really shows is that we have some mysterious faculty of "grasping concepts" (or "intuiting mathematical objects") and it is *this* that enables us to fix a model as *the* model, and not just operational and theoretical constraints; but this appeal to mysterious faculties seems both unhelpful as epistemology and unpersuasive as science. What neural process, after all, could be described as the perception of a mathematical object? Why of *one* mathematical object rather than another? I do not doubt that *some* mathematical axioms are built in to our notion of rationality ("every number has a successor"); but, if the axiom of choice and the continuum hypothesis are not, then, I am suggesting, Skolem's argument, or the foregoing extension of it, casts doubt on the view that these statements have a truth value independent of the theory in which they are embedded.

Now, suppose this is right and the axiom of choice is true when taken in the sense that it receives from *our* embedding theory and false when taken in the sense that it receives from extra-terrestrial theory. Urging this relativism is not advocating *unbridled* relativism; I do not doubt that there are some objective (if evolving) canons of rationality; I simply doubt that we would regard them as settling this sort of question, let alone as singling out *one* unique "rationally acceptable set theory". If this is right, then one is inclined to say that the extra-terrestrials have decided to let the axiom of choice be false and we have decided to let it be true; or that we have different "conventions"; but, of course, none of these words is literally right. It may well be the case that the idea that statements have their truth values *independent* of embedding theory is so deeply built into our ways of talking that there is simply no "ordinary language" word or short phrase which refers to the theory-dependence of meaning and truth. Perhaps this is why Poincaré was driven to exclaim "Convention, yes! Arbitrary, no!" when he was trying to express a similar idea in another context.

Is the problem a problem with the notion of a "set"? It would be natural to suppose that the problem Skolem points out, the problem of a surprising "relativity" of our notions, has to do with the notion of a "set", given the various problems which are *known* to surround *that* notion, or, at least, has to do with the problem of reference to "mathematical objects". But this is not so.

To see why it is not so, let us consider briefly the vexed problem of reference to theoretical entities in physical science. Although this may seem to be a problem more for philosophers of science or philosophers of language than for logicians, it is a problem whose logical aspects have frequently been of interest to logicians, as is witnessed by the expressions "Ramsey sentence", "Craig translation", etc. Here again, the realist—or, at least, the hard-core metaphysical realist—wishes it to be the case that *truth* and *rational acceptability* should be *independent* notions. He wishes it to be the case that what, e.g., electrons *are* should be distinct (and possibly different from) what we believe them to be or even what we would believe them to be given the best experiments and the epistemically best theory. Once again, the realist—the hard-core metaphysical realist—holds that our intentions single out "the" model, and that our beliefs are then either true or false in "the" model *whether we can find out their truth values or not.*

To see the bearing of the Löwenheim-Skolem Theorem (or of the intimately related Gödel Completeness Theorem and its model-theoretic generalizations) on this problem, let us again do a bit of model construction. This time the operational constraints have to be handled a little more delicately, since we have need to distinguish operational concepts (concepts that describe what we see, feel, hear, etc., as we perform various experiments, and also concepts that describe our acts of picking up, pushing, pulling, twisting, looking at, sniffing, listening to, etc.) from nonoperational concepts.

To describe our operational constraints we shall need three things. First, we shall have to fix a sufficiently large "observational vocabulary". Like the "observational vocabulary" of the logical empiricists, we will want to include in this set—call it the set of "0-terms"—such words as "red", "touches", "hard", "push", "look at", etc. Second, we shall assume that there *exists* (whether we can define it or not) a set of S which can be taken to be the set of macroscopically observable things and events (observable with the human sensorium, that means). The notion of an observable thing or event is surely vague; so we shall want S to be a generous set, that is, God is to err in the direction of counting too many things and events as "observable for humans" when He defines the set S, if it is necessary to err in either direction, rather than to err in the direction of leaving out some things that might be counted as borderline "observables". If one is a realist, then such a set S must exist, of course, even if our knowledge of the world and the human sensorium does not permit *us* to define it at the present time. The reason we allow S to contain events (and not just things) is that, as Richard Boyd has pointed out, some of the entities we can directly observe are *forces*—we can *feel* forces—and forces are not objects. But I assume that forces can be construed as predicates of either objects, e.g., our bodies, or of suitable events.

The third thing we shall assume given is a valuation (call it, once again 'OP') which assigns the correct truth value to each n-place 0-term (for $n = 1, 2, 3, \ldots$) on each n-tuple of elements of S on which it is defined. 0-terms are in general also defined on things not in S; for example, two molecules too small to see with the naked eye may touch, a dust-mote too small to see may be black, etc. Thus OP is a *partial* valuation in a double sense; it is defined on only a subset of the predicates of the language, namely the 0-terms, and even on these it only fixes a part of the extension, namely the extension of $T \restriction S$ (the restriction of T to S), for each 0-term T.

Once again, it is the valuation OP that captures our "operational constraints". Indeed, it captures these "from above", since it may well contain *more* information than we could actually get by using our bodies and our senses in the world.

What shall we do about "theoretical constraints"? Let us assume that there exists a possible formalization of present-day total science, call it 'T', and also that there exists a possible formalization of *ideal* scientific theory, call it 'T_I'. T_I is to be "ideal" in the sense of being *epistemically* ideal *for humans*. Ideality, in this sense, is a rather vague notion; but we shall assume that, when God makes up T_I, He constructs a theory which it would be rational for scientists to accept, or which is a limit of theories that it would be rational to accept, as more and more evidence accumulates, and also that he makes up a theory which is compatible with the valuation OP.

Now, 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 predictions, and thus conflict with OP. But T_I, by hypothesis, does not lead to any false predictions. Still, the metaphysical realist claims—and it is just this claim that makes him a *metaphysical* as opposed to an empirical realist—that T_I may be, in reality, false. What is not knowable as true may nonetheless be true; what is epistemically most justifiable to believe may nonetheless be false, on this kind of realist view. The striking connection between issues and debates in the philosophy of science and issues and debates in the philosophy of mathematics is that this sort of realism runs into *precisely* the same difficulties that we saw Platonism run into. Let us pause to verify this.

Since the ideal theory T_I must, whatever other properties it may or may not have, have the property of being *consistent*, it follows from the Gödel Completeness Theorem (whose proof, as all logicians know, is intimately related to one of Skolem's proofs of the Löwenheim-Skolem Theorem), that T_I has models. We shall assume that T_I contains a primitive or defined term denoting each member of S, the set of "observable things and events". The assumption that we made, that T_I agrees with OP, means that all those sentences about members of S which OP requires to be true are theorems of T_I. Thus if M is any model of T_I, M has to have a member corresponding to each member of S. We can even replace each member of M which corresponds to a member of S by that member of S itself, modifying the interpretation of the predicate letters accordingly, and obtain a model M' in which each term denoting a member of S in the "intended" interpretation does denote that member of S. Then the extension of each 0-term in that model will be partially correct to the extent determined by OP: that is, everything that OP "says" is in the extension of $\underline{P}$ is in the extension of $\underline{P}$, and everything that OP "says" is in the extension of the complement of $\underline{P}$ is in the extension of the complement of $\underline{P}$, for each 0-term, in any such model. In short, such a model is standard with respect to $P \upharpoonright S$ (P restricted to S) for each 0-term P.

Now, such a model satisfies all operational constraints, since it agrees with OP. It satisfies those theoretical constraints we would impose in the ideal limit of inquiry. So, once again, it looks as if any such model is "intended"—for what else could single out a model as "intended" than this? But if this is what it *is* to be an "intended model", T_I must be *true*—true in all intended models! The metaphysical realist's claim that even the ideal theory T_I might be false "in reality" seems to collapse into unintelligibility.

Of course, it might be contended that "true" does not follow from "true in all intended models". But "true" is the same as "true in the intended *interpretation*" (or "in *all* intended interpretations", if there may be more than one interpretation intended—or permitted—by the speaker), on any view. So to follow this line—which is, indeed, the right one, in my view—one needs to develop a theory on which interpretations are specified *other* than my specifying models.

Once again, an appeal to mysterious powers of the mind is made by some. Chisholm (following the tradition of Brentano) contends that the mind has a faculty of *referring to external objects* (or perhaps to external properties) which he calls by the good old name "intentionality". Once again most naturalistically minded philosophers (and, of course, psychologists), find the postulation of unexplained mental faculties unhelpful epistemology and almost certainly bad science as well.

There are two main tendencies in the philosophy of science (I hesitate to call them "views", because each tendency is represented by many different detailed views) about the way in which the reference of theoretical terms gets fixed. According to one tendency, which we may call the Ramsey tendency, and whose various versions constituted the received view for many years, theoretical terms come in batches or clumps. Each clump—for example, the clump consisting of the primitives of electromagnetic theory—is defined by a theory, in the sense that all the models of that theory which are standard on the observation terms count as intended models. The theory is "true" just in case it has such a model. (The "Ramsey sentence" of the theory is just the second-order sentence that asserts the existence of such a model.) A sophisticated version of this view, which amounts to relativizing the Ramsey sentence to an open set of "intended applications", has recently been advanced by Joseph Sneed.

The other tendency is the realist tendency. While realists differ among themselves even more than proponents of the (former) received view do, realists unite in agreeing that a theory may have a true Ramsey sentence and not be (in reality) true.

The first of the two tendencies I described, the Ramsey tendency, represented in the United States by the school of Rudolf Carnap, accepted the "relativity of theoretical notions", and abandoned the realist intuitions. The second tendency is more complex. Its, so to speak, conservative wing, represented by Chisholm, joins Plato and the ancients in postulating mysterious powers wherewith the mind "grasps" concepts, as we have already said. If we have more available with which to fix the intended model than merely theoretical and operational constraints, then the problem disappears. The radical pragmatist wing, represented, perhaps, by Quine, is willing to give up the intuition that T_I might be false "in reality". This radical wing is "realist" in the sense of being willing to assert that present-day science, taken more or less at face value (i.e., without philosophical reinterpretation) is at least approximately true; "realist" in the sense of regarding reference as trans-theoretic (a theory with a true Ramsey sentence may be false, because later inquiry may establish an incompatible theory as better); but not *metaphysical* realist. It is the moderate "center" of the realist tendency, the center that would like to hold on to metaphysical realism *without* postulating mysterious powers of the mind that is once again in deep trouble.

Pushing the problem back: the Skolemization of absolutely everything. We have seen that issues in the philosophy of science having to do with reference of theoretical terms and issues in the philosophy of mathematics having to do with the problem of singling out a unique "intended model" for set theory are both connected with the Löwenheim-Skolem Theorem and its near relative, the Gödel Completeness Theorem. Issues having to do with reference also arise in philosophy in connection with sense data and material objects and, once again, these connect with the model-theoretic problems we have been discussing. (In some way, it really seems that the Skolem Paradox underlies the *characteristic* problems of 20th century philosophy.)

Although the philosopher John Austin and the psychologist Fred Skinner both tried to drive sense data out of existence, it seems to me that most philosophers and psychologists think that there are such things as *sensations*, or *qualia*. They may not be objects of perception, as was once thought (it is becoming increasingly fashionable to view them as states or conditions of the sentient subject, as Reichenbach long ago urged we should); we may not have incorrigible knowledge concerning them; they may be somewhat ill-defined entities rather than the perfectly sharp particulars they were once taken to be; but it seems reasonable to hold that they are part of the legitimate subject matter of cognitive psychology and philosophy and not mere pseudo-entities invented by bad psychology and bad philosophy.

Accepting this, and taking the operational constraint this time to be that we wish the ideal theory to correctly predict all sense data, it is easily seen that the previous argument can be repeated here, this time to show that (if the "intended" models are the ones which satisfy the operational and theoretical constraints we now have, or even the operational and theoretical constraints we would impose in some limit) then, either the present theory is "true", in the sense of being "true in all intended models", provided it leads to no false predictions about sense data, or else the ideal theory is "true". The first alternative corresponds to taking the theoretical constraints to be represented by current theory; the second alternative corresponds to taking the theoretical constraints to be represented by the ideal theory. This time, however, it will be the case that even terms referring to ordinary material objects—terms like 'cat' and 'dog'—get differently interpreted in the different "intended" models. It seems, this time, as if we cannot even refer to ordinary middle sized physical objects except as formal constructs variously interpreted in various models.

Moreover, if we agree with Wittgenstein that the *similarity relation* between sense data we have at different times is not itself something present to my mind—that "fixing one's attention" on a sense datum and thinking "by 'red' I mean whatever is like *this*" does not really pick out any relation of similarity at all—and make the natural move of supposing that the intended models of my language when I now and in the future talk of the sense data I had at some past time t_0 are singled out by operational and theoretical constraints, then, again, it will turn out that my *past* sense data are mere formal constructs which get differently interpreted in various models. If we further agree with Wittgenstein that the notion of truth requires a *public* language (or requires at least states of the self at more than one time —that a "private language for one specious present" makes no sense), then even my *present* sense data are in this same boat . . . . In short, one can "Skolemize" absolutely everything. It seems to be absolutely impossible to fix a determinate reference (without appeal to nonnatural mental powers) for *any* term at all. If we apply the argument to the very metalanguage we use to talk about the predicament . . . ?

The same problem has even surfaced recently in the field of cognitive psychology. The standard model for the brain/mind in this field is the modern computing machine. This computing machine is thought of as having something analogous to a formalized language in which it computes. (This hypothetical brain language has even received a name—"mentalese".) What makes the model of cognitive psychology a *cognitive* model is that "mentalese" is thought to be a medium whereby the brain constructs an *internal representation* of the external world. This idea runs immediately into the following problem: if "mentalese" is to be a vehicle for describing the external world, then the various predicate letters must have extensions which are sets of external things (or sets of *n*-tuples of external things). But if the way "mentalese" is "understood" by the deep structures in the brain that compute, record, etc. in this "language" is *via* what artificial intelligence people call "procedural semantics"—that is, if the brain's *program for using* "mentalese" comprises its entire "understanding" of "mentalese"—where the program for using "mentalese", like any program, refers only to what is *inside* the computer—then how do *extensions* ever come into the picture at all? In the terminology I have been employing in this address, the problem is this: if the extension of predicates in "mentalese" is fixed by the theoretical and operational constraints "hard wired in" to the brain, or even by theoretical and operational constraints that it evolves in the course of inquiry, then these will not fix a *determinate* extension for any predicate. If thinking is ultimately done in "mentalese", then *no concept we have will have a determinate extension*. Or so it seems.

The bearing of causal theories of reference. The term "causal theory of reference" was originally applied to my theory of the reference of natural kind terms and Kripke's theory of the reference of proper names. These theories did not attempt to *define* reference, but rather attempted to say something about how reference is fixed, if it is not fixed by associating definite descriptions with the terms and names in question. Kripke and I argued that the intention to preserve reference through a historical chain of uses and the intention to cooperate socially in the fixing of reference make it possible to use terms successfully to refer although no one definite description is associated with any term by all speakers who use that term. These theories assume that individuals can be singled out for the purpose of a "naming ceremony" and that inferences to the existence of definite theoretical entities (to which names can then be attached) can be successfully made. Thus these theories did not address the question as to how any term can acquire a determinate reference (or any gesture, e.g., pointing—of course, the "reference" of gestures is just as problematic as the reference of terms, if not more so). Recently, however, it has been suggested by various authors that some account can be given of how at least some basic sorts of terms refer in terms of the notion of a "causal chain". In one version,[5] a version strikingly reminiscent of the theories of Ockham and other 14th century logicians, it is held that a term refers to "the dominant source" of the beliefs that contain the term. Assuming we can circumvent the problem that the dominant cause of our beliefs concerning *electrons* may well be *textbooks*,[6] it is important to notice that even if a *correct* view of this kind can be elaborated, it will do nothing to resolve the problem we have been discussing.

The problem is that adding to our hypothetical formalized language of science a body of theory titled "causal theory of reference" *is* just adding more *theory*. 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 *all true sentences*, 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 metalanguage 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, *in each model M*, $reference_M$ is definable in terms of $causes_M$; but, unless the word 'causes' (or whatever the causal predicate or predicates may be) is already glued to one definite relation with metaphysical glue, this does not fix a determinate extension for 'refers' at all.

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 environment" 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 "transduce" information from the environment, and such a description might well be what the causal theorists are looking for. The program of realism in the philosophy of science—of *empirical* realism, not metaphysical realism—is to show that scientific theories can be regarded as better and better representations of an objective world with which we are interacting; if such a view is to be part of science itself, as empirical realists contend it should be, then the interactions with the world by means of which this representation is formed and modified 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.

Ideal theories and truth. One reaction to the problem I have posed would be to say: there are many ideal theories in the sense of theories which satisfy the operational constraints, and in addition have all the virtues (simplicity, coherence, con-

[5] Cf. Gareth Evans, *The causal theory of names*, ***Aristotelian Society Supplementary Volume*** XLVII, pp. 187–208, reprinted in ***Naming, necessity and natural kinds,*** (Stephen P. Schwartz, Editor), Cornell University Press, 1977.

[6] Evans handles this case by saying that there are appropriateness conditions on the type of causal chain which must exist between the item referred to and the speaker's body of information.

taining the axiom of choice, whatever) that humans like to demand. But there are no "facts of the matter" not reflected in constraints on ideal theories in this sense. Therefore, what is really true is what is common to all such ideal theories; what is really false is what they all deny; all other statements are neither true nor false.

Such a reaction would lead to too few truths, however. It may well be that there are rational beings—even rational human species—which do not employ our color predicates, or who do not employ the predicate "person", or who do not employ the predicate "earthquake".[7] I see no reason to conclude from this that *our* talk of red things, or of persons, or of earthquakes, lacks truth value. If there are many ideal theories (and if "ideal" is itself a somewhat interest-relative notion), if there are many theories which (given appropriate circumstances) it is perfectly rational to accept, then it seems better to say that, insofar as these theories say different (and sometimes, apparently incompatible) things, that some facts are "soft" in the sense of depending for their truth value on the speaker, the circumstances of utterance, etc. This is what we have to say in any case about cases of ordinary vagueness, about ordinary causal talk, etc. It is what we say about apparently incompatible statements of simultaneity in the special theory of relativity. To grant that there is more than one true version of reality is not to deny that some versions are false.

It may be, of course, that there *are* some truths that *any* species of rational inquirers would eventually acknowledge. (On the other hand, the set of these may be empty, or almost empty.) But to say that *by definition* these are all the truths there are is to redefine the notion in a highly restrictive way. (It also assumes that the notion of an "ideal theory" is perfectly clear; an assumption which seems plainly false.)

Intuitionism. It is a striking fact that this entire problem does *not* arise for the standpoint of mathematical intuitionism. This would not be a surprise to Skolem: it was precisely his conclusion that "most mathematicians want mathematics to deal, ultimately, with performable computing operations and not to consist of formal propositions about objects called this or that."

In intuitionism, knowing the meaning of a sentence or predicate consists in associating the sentence or predicate with a procedure which enables one to recognize when one has a proof that the sentence is constructively true (i.e., that it is possible to carry out the constructions that the sentence asserts can be carried out) or that the predicate applies to a certain entity (i.e., that a certain full sentence of the predicate is constructively true). The most striking thing about this standpoint is that the *classical notion of truth is nowhere used*—the semantics is entirely given in terms of the notion of "constructive proof", *including the semantics of "constructive proof" itself.*

Of course, the intuitionists do not think that "constructive proof' can be formalized, or that "mental constructions" can be identified with operations in our *brains*. Generally, they assume a strongly intentionalist and *a prioristic* posture in philosophy—that is, they assume the existence of mental entities called "meanings" and of a special faculty of intuiting constructive relations between these entities.

[7] For a discussion of this very point, cf. David Wiggins, ***Truth, invention and the meaning of life,*** British Academy, 1978.

These are not the aspects of intuitionism I shall be concerned with. Rather I wish to look on intuitionism as an example of what Michael Dummett has called "nonrealist semantics"—that is, 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.

The problem with realist semantics—truth-conditional semantics—as Dummett has emphasized, is that if we hold that the understanding of the sentences of, say, set theory consists in our knowledge of their "truth conditions", then how can we possibly say what *that* knowledge in turn consists in? (It cannot, as we have just seen, consist in the use of language or "mentalese" under the control of operational plus theoretical constraints, be they fixed or evolving, since such constraints are too weak to provide a determinate extension for the terms, and it is this that the realist wants.)

If, however, the understanding of the sentences of a mathematical theory consists in the mastery of verification procedures (which need not be fixed once and for all—we can allow a certain amount of "creativity"), then a mathematical theory can be completely understood, and this understanding does not presuppose the notion of a "model" at all, let alone an "intended model".

Nor does the intuitionist (or, more generally, the "nonrealist" semanticist) have to foreswear *forever* the notion of a model. He has to foreswear reference to models in his account of *understanding*; but, once he has succeeded in understanding a rich enough language to serve as a metalanguage for some theory T (which may itself be simply a sublanguage of the metalanguage, in the familiar way), he can define 'true in T' à la Tarski, he can talk about "models" for T, etc. He can even define 'reference' or ('satisfaction') exactly as Tarski did.

Does the whole "Skolem Paradox" arise again to plague him at this stage? The answer is that it does not. To see why it does not, one has to realize what the "existence of a model" means in *constructive* mathematics.

"Objects" in constructive mathematics are *given through descriptions*. Those descriptions do not have to be mysteriously attached to those objects by some nonnatural process (or by metaphysical glue). Rather the possibility 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 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.

Intuitionism liberalized. It is not my aim, however, to try to convert my audience to intuitionism. Set theory may not be the "paradise" Cantor thought it was, but it is not such a bad neighborhood that I want to leave of my own accord, either. Can we separate the philosophical idea behind intuitionism, the idea of "nonrealist" semantics, from the restrictions and prohibitions that the historic intuitionists wished to impose upon mathematics?

The answer is that we can. First, as to set theory: the objection to *impredicativity*, which is the intuitionist ground for rejecting much of classical set theory, has little or no connection with the insistence upon verificationism itself. Indeed, intuitionist mathematics is itself "impredicative", inasmuch as the intuitionist notion of constructive proof presupposes constructive proofs which refer to the totality of *all* constructive proofs.

Second, as to the propositional calculus: it is well known that the classical connectives can be reintroduced into an intuitionist theory by reinterpretation. The important thing is not whether one uses "classical propositional calculus" or not, but how one *understands* the logic if one does use it. Using classical logic as an intuitionist would understand it, means, for example, keeping track of when a disjunction is selective (i.e., one of the disjuncts is constructively provable), and when it is nonselective; but this does not seem like too bad an idea.

In short, while intuitionism may go with a greater interest in constructive mathematics, a liberalized version of the intuitionist standpoint need not rule out "classical" mathematics as either illegitimate or unintelligible. What about the language of empirical science? Here there are greater difficulties. Intuitionist logic is given in terms of a notion of *proof*, and proof is supposed to be a *permanent* feature of statements. Moreover, proof is nonholistic; there is such a thing as the proof (in either the classical or the constructive sense) of an isolated mathematical statement. But verification in empirical science is a matter of degree, not a "yes-or-no" affair; even if we made it a "yes-or-no" affair in some arbitrary way, verification is a property of empirical sentences that can be *lost*; in general the "unit of verification" in empirical science is the theory and not the isolated statement.

These difficulties show that sticking to the intuitionist standpoint, however liberalized, would be a bad idea in the context of formalizing empirical science. But they are not incompatible with "nonrealist" semantics. The crucial question is this: do we think of the *understanding* of the language as consisting in the fact that speakers possess (collectively if not individually) an evolving network of verification procedures, or as consisting in their possession of a set of "truth conditions"? If we choose the first alternative, the alternative of "nonrealist" semantics, then the "gap" between words and world, between our *use* of the language and its "objects", never appears.[8] Moreover, the "nonrealist" semantics is not *inconsistent* with

[8] To the suggestion that we identify truth with 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 scientific development differ in such a way to make $\bar{A}$ part of the ideal theory accepted in the long run; in that circumstance, 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 *sentence* $\ulcorner A \urcorner$ has a fixed meaning independent of what version we accept.

More deeply, as Michael Dummett first pointed out, 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 *understood 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.

realist semantics; it is simply *prior* to it, in the sense that it is the "nonrealist" semantics that must be internalized if the language is to be understood.

Even if it is not inconsistent with realist semantics, taking the nonrealist semantics as our picture of how the language is understood undoubtedly will affect the way we view questions about reality and truth. For one thing, verification in empirical science (and, to a lesser extent, in mathematics as well, perhaps) sometimes depends on what we before called "decision" or "convention". Thus facts may, on this picture, depend on our interests, saliencies and decisions. There will be many "soft facts". (Perhaps whether $V = L$ or not is a "soft fact".) I cannot, myself, regret this. If appearance and reality end up being endpoints on a continuum rather than being the two halves of a monster Dedekind cut in all we conceive and do not conceive, it seems to me that philosophy will be much better off. The search for the "furniture of the Universe" will have ended with the discovery that the Universe is not a furnished room.

Where did we go wrong?—The problem solved. What Skolem really pointed out is this: no interesting theory (in the sense of first-order theory) can, in and of itself, determine its own objects up to isomorphism. Skolem's argument can be extended as we saw, to show that if theoretical constraints do not determine reference, then the addition of operational constraints will not do it either. It is at this point that reference itself begins to seem "occult"; that it begins to seem that one cannot be any kind of a realist without being a believer in nonnatural mental powers. Many moves have been made in response to this predicament, as we noted above. Some have proposed that *second-order* formalizations are the solution, at least for mathematics; but the "intended" interpretation of the second-order formalism is not fixed by the use of the formalism (the formalism itself admits so-called "Henkin models", i.e., models in which the second-order variables fail to range over the *full* power set of the universe of individuals), and it becomes necessary to attribute to the mind special powers of "grasping second-order notions". Some have proposed to accept the conclusion that mathematical language is only partially interpreted, and likewise for the language we use to speak of "theoretical entities" in empirical science; but then are "ordinary material objects" any better off? Are sense data better off? Both Platonism and phenomenalism have run rampant at different times and in different places in response to this predicament.

The problem, however, lies with the predicament itself. The predicament only *is* a predicament because we did two things: first, we gave an account of understanding the language in terms of programs and procedures for *using* the language (what else?); then, secondly, we asked what the possible "models" for the language were, thinking of the models as existing "out there" *independent of any description.* At this point, something really weird had already happened, had we stopped to notice. On any view, the understanding of the language must determine the reference of the terms, or, rather, must determine the reference given the context of use. If the use, even in a fixed context, does not determine reference, then use is not understanding. The language, on the perspective we talked ourselves into, has a full program of use; but it still lacks an *interpretation.*

This is the fatal step. To adopt a theory of meaning according to which a language whose whole use is specified still lacks something—viz. its "interpretation"—is to accept a problem which *can* only have crazy solutions. To speak as if *this* were my problem, "I know how to use my language, but, now, how shall I single out an interpretation?" is to speak nonsense. Either the use *already* fixes the "interpretation" or *nothing* can.

Nor do "causal theories of reference", etc., help. Basically, trying to get out of this predicament by *these* means is hoping that the *world* will pick one definite extension for each of our terms even if *we* cannot. But the world does not pick models or interpret languages. *We* interpret our languages or nothing does.

We need, therefore, a standpoint which links use and reference in just the way that the metaphysical realist standpoint refuses to do. The standpoint of "nonrealist semantics" is precisely that standpoint. From that standpoint, it is trivial to say that a model in which, as it might be, the set of cats and the set of dogs are permuted (i.e., 'cat' is assigned the set of dogs as its extension, and 'dog' is assigned the set of cats) is "unintended" even if corresponding adjustments in the extensions of all the other predicates make it end up that the operational and theoretical constraints of total science or total belief are all "preserved". Such a model would be unintended *because we do not intend the word 'cat' to refer to dogs*. From the metaphysical 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 *its* intended interpretation singled out; but we are in the same predicament with respect to the metalanguage that we are in with respect to the object language, from that standpoint, so all is in vain. However, from the viewpoint of "nonrealist" semantics, the metalanguage is completely understood, and so is the object language. So we can *say and understand*, " 'cat' refers to cats". Even though the model referred to satisfies the theory, etc., it is "unintended"; we recognize that it is unintended *from the description through which it is given* (as in the intuitionist case). Models are not lost noumenal waifs looking for someone to name them; they are constructions within our theory itself, and they have names from birth.

DEPARTMENT OF PHILOSOPHY
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS 02138