Semantics without Reference

CHRISTOPHER GAUKER*

Abstract A theory of reference may be either an *analysis* of reference or merely an account of the correct use of the verb "to refer". If we define the validity of arguments in the standard way, in terms of assignments of individuals and sets to the nonlogical vocabulary of the language, then, it is argued, we will be committed to seeking an *analysis* of reference. Those who prefer a metalinguistic account, therefore, will desire an alternative to standard semantics. One alternative is the Quinean conception of validity as essentially a matter of logical form. Another alternative is Leblanc's truth-value semantics. But these prove to be either inadequate for purposes of metatheory or philosophically unsatisfactory. This paper shows how validity (i.e., semantic consequence) may be defined in a way that avoids the problems facing these other alternatives to standard semantics and also permits a metalinguistic account of reference. The validity of arguments is treated as a matter of logical form, but validity for forms is defined on analogy with the definition of semantic consequence in truth-value semantics.

The contemporary concept of reference is tightly bound up with standard formal semantics. These entanglements constrain what we may accept as a theory of reference. Under these constraints, an acceptable theory of reference has proven very difficult to devise. The problem of reference might be easier if logic could make do with some other kind of semantics. One alternative might be to define validity in terms of logical form. Another might be to give a substitutional interpretation of the quantifiers. Unfortunately, both of these alternatives have

problems of their own. The purpose of this paper is to show how these two alternatives may be combined in a way that avoids the pitfalls of each and also releases the problem of reference from the constraints of standard semantics.

1 The problem   What is reference? Offhand, it appears to be a *relation*, just as *being heavier than* is a relation. Moreover, it is a relation that holds between *words* and other things. It holds between "chair" and the chair I'm sitting in, between "chair" and the chair next door, and between "chair" and the chairs of the past and future. It holds between "Socrates" and Socrates, between "meson" and mesons, and between "beautiful" and beautiful things. Of course, the reason it holds between these things is not just that it holds between every word and *everything*. For instance, it doesn't hold between "basketball" and daffodils. The fact that the relation *being heavier than* holds between an atom of oxygen and an atom of hydrogen, as well as between Pavarotti and Diana Ross, but not between everything and everything, ought to make us wonder what *being heavier than* amounts to, if we don't already know. Likewise, the combination of diversity and specificity exhibited by the reference relation ought to make us wonder what reference is. We mustn't just call it an "unanalyzable primitive".

Accounts of reference that might satisfy us fall into two classes. On the one hand, we might seek what I'll call an *analysis* of reference. An analysis takes the form:

$$t \text{ refers to } a \text{ if and only if } \ldots.$$

(The "if and only if" is assumed to be in some way stronger than the material biconditional. To some extent it ought to support counterfactual inferences.) On the other hand, we might satisfy ourselves with what I'll call a *metalinguistic* account of reference. A metalinguistic account of reference is one that *merely* explains the correct *use* of "refers". In other words, a metalinguistic account satisfies the following two conditions: (i) it provides a biconditional of the form:

$$\text{A sentence of the form } \ulcorner \tau \text{ refers to } \alpha \urcorner$$
$$\text{is correctly used if and only if } \ldots.$$

and similarly explains the correct use of other sorts of sentences containing forms of the verb "to refer"; and (ii) it does all this without implying any *analysis* of reference.

There are many ways to try to analyze reference. One way, represented by, for instance, Field [7] and Devitt [5], is to try to explain reference in nonsemantical terms. Such an analysis might involve causal relations, covariances, reference-passings, and so on. This strategy will be favored by those who wish to explain other semantic concepts, such as truth and meaning, in terms of reference. Other philosophers, for instance Davidson [4] and McDowell [19], are willing to take for granted that we understand *truth* in order to explain reference. Thus reference might be conceived as a theoretical relation postulated in the course of giving a theory of truth from which we may derive independently confirmable statements of the form $\ulcorner "s" \text{ is true in language } L \text{ if and only if } p \urcorner$. Yet another approach (attributed to David Kaplan) is to take *meaning* as given,

where meaning is conceived as something that may differentiate co-referring expressions, and then explain reference as a function of meaning and context. Thus, "I", when *I* say it, refers to *me* because of what it means and the fact that I'm the one who said it. Or again, the meaning of "water" is such that on Earth it refers to H_2O but on Twin-Earth it refers to XYZ.

Metalinguistic accounts of reference have been far less common but are not unheard of. The theories of Sellars [25], Brandom [2], Hartman [9], and Hill [10] may all be viewed as, in effect, reaching for a metalinguistic account. Moreover, metalinguistic approaches to the cognate concept *truth* may also be found in Ramsey [24] and Grover, Camp and Belnap [8]. To get a sense for what a metalinguistic theory of reference might look like, consider the following four inference rules:

(I) "*t*" refers to *a*
"*b*" is a translation of "*t*" into our language
———————————————————————
$a = b$.

(II) $a = b$
"*b*" is a translation of "*t*" into our language
———————————————————————
"*t*" refers to *a*.

(III) "*P*" refers to *F*-things
"*G*" is a translation of "*P*" into our language
———————————————————————
All and only *F*-things are *G*-things.

(IV) All and only *F*-things are *G*-things
"*G*" is a translation of "*P*" into our language
———————————————————————
"*P*" refers to *F*-things.

One representative but plainly inadequate metalinguistic account of reference would be that a sentence of the form $\ulcorner \tau$ refers to $\alpha \urcorner$ is correctly used if and only if it occurs as a premise or a conclusion in an inference having the form (I), (II), (III), or (IV). This account is representative in that it has the right form and does not imply an *analysis* of reference.[1] Of course, this account is inadequate since forms of the verb "to refer" may be used correctly in contexts other than such inferences as these. However, the hope behind the metalinguistic approach is that the meaning of "refers" is exhausted by its logical ties to other bits of language and that these ties can be exhaustively described.

In this paper I am going to assume that we ought to seek a metalinguistic account of reference rather than an analysis of reference. I am not going to try to justify that assumption, but if I were going to try to justify it, my rationale would be that none of the above-mentioned strategies for analysis seems to be very promising. This paper is for those who share my despair. However, my objective in this paper is not to try to develop a satisfactory metalinguistic account of reference but only to clear away a certain obstacle. The obstacle is standard formal semantics, which, as we shall see, is entangled with the concept of reference in a way that forces us to seek the sort of analysis of reference we will not have if we merely explain the correct use of "refers"

Here's what I mean by *standard semantics*. In standard semantics, an *interpretation* $\langle D, I_D \rangle$ of a first-order language L consists of an assignment I_D of elements of a nonempty *domain D* to the individual constants of L and of ordered *n*-tuples of elements of D to the *n*-ary predicates of L. Standard semantics gives a recursive definition of *truth-in-L* on $\langle D, I_D \rangle$ (or first a recursive definition of *satisfaction-in-L* on $\langle D, I_D \rangle$ by a denumerable sequence of members of D and then a definition of truth-in-L in terms of satisfaction). $\langle D, I_D \rangle$ is said to be a *model* for those formulas of L that are true-in-L on $\langle D, I_D \rangle$. Finally, standard semantics defines an argument in L to be *valid* just in case every model for the premises is a model for the conclusion as well.

There is a very great difference between the concept of a standard semantical *assignment* (belonging to an *interpretation*) and the concept of *reference*, which we desire an account of. This difference needs to be emphasized since some logicians call these assignments "reference relations" or something similar. Not every standard semantical assignment is really a reference relation. For instance, the assignment that assigns the set of gorillas to the English predicate "yellow" is not. And there is no problem about defining the *set* of these assignments. Moreover, reference, as usually conceived in contemporary philosophy of language, is not even one *particular* standard semantical assignment. It is a relation which is not restricted to any given language and which even the words of some not-yet-invented language may come to bear to things.

Nonetheless, the contemporary concept of reference is *bound up* with the concept of a standard semantical assignment. One tends to suppose that the reference relation, as it pertains to a given formal language, may be *represented* as a particular assignment in a standard semantical interpretation of the language. Further, formal semantics for the artificial languages of logic is supposed to shed light on the semantics of natural languages. Thus we imagine that we could define *interpretations* for natural languages roughly analogous to those that standard semantics defines for artifical languages, and define such concepts as validity and consistency for natural languages in terms of these interpretations. So we imagine that the reference relation as it pertains to a given natural language may also be represented as a particular standard semantical assignment in such an interpretation. Among the interpretations for a given language we imagine that there will be one special one, call it *the intended interpretation*, which will contain the assignment representing the reference relation for the language. The domain for the intended interpretation will be the set of all individuals in reality. The assignment given by the intended interpretation will assign to each nonlogical constant its referent, that is, what it really refers to. And in that language, truth on the intended interpretation will be truth *simpliciter*. Thus reality and reference, we suppose, *determine* a model for the sentences of the language that really are true. It is hard to see how we could think of standard semantics as the right way to do semantics and not think of the concept of reference as bound up with standard semantics in this way.[2]

Thus, in adopting standard semantics we are committed to a certain view of reference. Further, in committing ourselves to this view of reference, we appear to be committed to seeking an analysis of reference rather than a metalinguistic account of reference (assuming that we are committed to seeking *some* account of reference). The reason is that if reference, as it pertains to a given

language, is conceived as a standard semantical assignment of individuals to individual constants and of sets to predicates, then an account of reference must tell us *which* such assignment represents the reference relation for the language in question, i.e., which assignment belongs to the intended interpretation, and it seems that any such account will amount to an analysis of the relation that obtains between a nonlogical constant and the individual or set that the privileged assignment assigns to it.

Of course, we should not expect an account of reference to identify the reference relation in the sense of explicitly telling us which individuals belong to the extensions of all of our predicates and which individuals our names refer to. If our theory of reference did that, then we could use it as a substitute for empirical research. To decide whether any given sentence was true, we could simply check the referents of its constituent terms. Since we should not expect a theory of reference to be a substitute for empirical research, we should not expect to identify the reference relation in *that* sense.

Still, we might expect a theory of reference to provide a description that only the reference relation would satisfy, and in *this* sense it might tell us which standard semantical assignment represents the reference relation as it pertains to the language in question. My claim is that if reference is thus conceived as a standard semantical assignment then a theory of reference must in this sense tell us which standard semantical assignment represents the reference relation and that only an analysis of reference, not a metalinguistic account of reference, will do that. But perhaps it is not obvious that only an analysis would do, and so now I want to try to articulate my sense that a metalinguistic account would not.

First, consider that sentences of the form $\ulcorner \tau$ refers to $\alpha \urcorner$ are sometimes *true (simpliciter)*. On the view of reference according to which reference is entangled with standard semantics in the manner described above, what that means is this: τ refers to some expression t, α refers to some thing a (an individual or set), and $\langle t, a \rangle$ belongs to the set that "refers to" refers to, namely the set of ordered pairs $\langle x, y \rangle$ such that x refers to y. For example, " 'Robots' refers to robots" is true because " 'Robots' " refers to "robots", "robots" refers to the set of robots, and the pair $\langle$"robots", the set of robots$\rangle$ belongs to the set of ordered pairs that "refers to" refers to. Thus, if we adopt a standard semantics, the question that an account of reference has to answer is, what qualifies an ordered pair for membership in the set of ordered pairs $\langle x, y \rangle$ such that x refers to y? However, this question would not be answered by a metalinguistic account of reference. That is, this question would not be answered merely by explaining the correct use of "refers", as in sentences of the form $\ulcorner \tau$ refers to $\alpha \urcorner$. These explanations would be about τ and α, i.e., the *words* that we may place before and after "refers to" (normally the quotation-name of a term and the name of an individual or set). However, what we have to explain is something about t and a, i.e., the things that reference relates (normally a term and the individual or set it refers to). In order to explain in a general way the conditions under which an ordered pair belongs to the set of ordered pairs $\langle x, y \rangle$ such that x refers to y, we will need to spell out in a general way the conditions under which x refers to y, and that means that we will need an *analysis* of reference in the sense I have defined.

I am not sure that everybody will be persuaded by this argument because it

turns on an application of the conception of reference entangled with standard semantics to the special case of the predicate "refers to" and because it speaks of what the terms τ and α are *about*. So now I'll try a second, independent, argument to show that in viewing reference as bound up with standard semantics we are committed to seeking an analysis of reference rather than a metalinguistic account. Again, the question is whether a metalinguistic account of reference could succeed in specifying *which* of all possible assignments of objects and sets to terms of L is the reference relation for L. The answer, I am going to argue, is "not in general".

To begin, I think we can all agree that we do not discover truths by inspecting reference relations (or at least this is not the main way). So we may rule out some assignments by confining the reference relation, as it pertains to L, to those assignments of objects and sets to terms of L on which the sentences of L that we independently take to be true turn out true. In short, we may confine our attention to those assignments that belong to interpretations that are models for our theory of the world as formulated in L. However, any set of formulas that has at least one model is bound to have more than one. (For instance, other models can be got from a first model by a one-to-one substitution of elements of the domain for elements of the domain.) So more than one assignment is going to satisfy this condition. If this were the only condition the reference relation had to satisfy, reference would be wildly and inevitably indeterminate. So there must be more to a theory of reference than this.[3]

Thus our question becomes whether an account of the correct use of "refers" would suffice to isolate the reference relation, as it pertains to L, from all the "unintended" models of our theory of the world as formulated in L. In order to address this question, we need to distinguish between the primitive vocabulary of L and what we might call the "primitive referrers" of L. Roughly, the primitive vocabulary of L consists of those vocabulary items of L that cannot be generated (either alone or as part of some longer expression) by applying the formation rules of L. The primitive referrers of L, on the other hand, are those vocabulary items of L to which a standard semantical assignment makes some assignment (not on the basis of some recursion). The primitive referrers need not in general be included in the primitive vocabulary. For instance, we might have a language in which an infinite number of individual constants were generated by concatenating the letter "a" with one or more apostrophes.

If there are only finitely many primitive referrers in L, then perhaps a purely metalinguistic account of reference will suffice to isolate the intended interpretation (leaving aside my first argument above). If there are only finitely many primitive referrers to which referents need be assigned, then a specification of the reference relation as it pertains to L may take the form of a mere *list* of sentences of the form $\ulcorner \tau$ refers to $\alpha \urcorner$, and to decide whether any such sentence belonged on our list we could perhaps apply our metalinguistic account of reference, i.e., our account of the conditions under which sentences of that form are correctly used. On the other hand, if the number of primitive referrers in L is infinite, then I doubt whether a metalinguistic account will suffice for specification of the intended interpretation. Granted, by *applying* our account we might rule out *some* of the possible assignments. Whenever we find by that ac-

count that we are licensed to assert that t refers to a, we may rule out all assignments in which something other than a is assigned to t. But if the number of primitive referrers in L is infinite, such a procedure will never isolate a unique assignment. To identify the reference relation for L we would need to say in a general way what it takes for a term t of L to refer to an individual or set a, and doing that would amount to giving an analysis of reference (as it pertains to L).

So the question whether we can have both standard semantics and a metalinguistic account of reference becomes the question whether we need to countenance any language containing an infinite number of primitive referrers. Perhaps there is no real need to countenance languages having an infinite primitive *vocabulary*, and maybe there would be something fundamentally unlearnable about such languages. But if we think of reference as a standard semantical assignment of a privileged sort, then it is hard to see how we can escape from even *considering* languages containing an infinite number of primitive referrers. Consider languages containing a name for each natural number (formed, of course, from a finite vocabulary). Or suppose that the best theory of adverbs does *not* treat the reference of expressions of the form VERB + ADVERBIAL-PHRASE as an explicable function of the reference of VERB and the reference of ADVERBIAL-PHRASE, or that the reference of verb phrases of the form ⌜believes that p⌝ cannot be treated as an explicable function of the reference of component parts drawn from a finite stock. Well, surely we have to countenance languages containing infinitely many adverbial phrases and infinitely many verb phrases of the form ⌜believes that p⌝. So in these cases we would have to allow for an infinite number of primitive referrers. So if we want to hold on to standard semantics and be able to contemplate these theoretical contingencies, we will have to hold out for an *analysis* of reference as opposed to a metalinguistic account.

The upshot is that if we think that standard semantics is the right way to do formal semantics then we will inevitably think of reference as bound up with standard semantics in the ways I have described and so will be unable to explicate reference merely by explaining the correct use of "refers". So if this is how we want to explain reference, which it is, we will need an alternative to standard semantics. In the next section I will describe two alternatives that ultimately fail to satisfy our needs. Then in Section 3 I will set out my own alternative.

In philosophizing about logic there is always the problem of how to relate the sentences of our artificial languages to the sentences of natural language, which are what ultimately interest us. Some philosophers deal with this by pretending that the artificial languages are fragments of a natural language. My way will be to suppose that the sentences of our artificial languages have translations into English. I think this makes sense even in a context such as this, where the nature of reference is the issue. In any case, I think that everything I say can be reformulated under the pretense that the artificial languages are fragments of English.

2 Alternatives   Elementary logic textbooks normally define validity for arguments in one of two ways. One way, which I'll call the *one-step* approach, is basically that of standard semantics. On this approach, an argument is defined as valid just in case the conclusion is true on every *interpretation* on which the

premises are all true. (A typical example of this approach is Bergmann, Moor and Nelson [1].) The other way, which probably stems from Quine in most cases, is what I'll call the *two-step* approach. Here an argument is said to be valid just in case it has a valid form (or schema). An argument *form*, in turn, is said to be valid just in case it has no counterexample. A counterexample, of course, is an argument having that form in which the premises are in fact true and the conclusion is in fact false. (A typical example of this approach is Klenk [12].)

In the context of propositional logic, the two-step approach is probably the most natural, pedagogically speaking. It affords a simple explanation of what we are "saying" with truth tables. The letters that head the columns, we may explain, are *schematic* letters, and each row of the table represents a class of substitutions for these schematic letters. For instance, the first row of a standard truth table represents the class of substitutions in which every sentence substituted for a schematic letter is *true*. On the one-step approach, by contrast, the letters that head the columns are *constants* and the rows, we have to explain, represent "assignments of truth values" to these constants. The pedagogical problem is that it is hard to make intuitive sense of the idea of assigning truth values without dipping into predicate logic (I'll return to this). In the context of predicate logic, however, it is just as easy to take the one-step approach and explain that the various interpretations of the nonlogical constants represent alternative denotations or extensions.

A virtue of the two-step approach, from my point of view, is precisely that it avoids all talk of *interpretations* and thus avoids those entanglements with the concept of reference that give us trouble. The problem with the two-step approach, however, is that without some finagling it is unsuitable for use in proving the completeness of the predicate calculus. If we take the two-step approach, then to prove (strong) completeness we have to show that every argument form of every underivable argument has a counterexample. Since there are infinitely many underivable arguments, it will not be possible actually to *produce* a counterexample to every form of every underivable argument. Consequently, we will be driven to more abstract methods. The usual method (stemming from Henkin) involves defining truth on an interpretation as in standard semantics. But if we are going to make use of interpretations in proving completeness, we might as well define validity directly in terms of them as in the one-step method of standard semantics. In any case, we will be stuck with the entanglements with reference.

Quine is a good example of someone who indulges in this kind of bad faith. According to him, the most basic explanation of the concept of logical truth (validity for sentences) is something like this: A logical truth is a sentence from which we get only truths when we substitute sentences (simple or compound, open or closed) for simple sentences, taking care to match variables appropriately ([23], Chapter 4, pp. 50,58). But when he goes to do things in a rigorous way, he defines logical truth like this: A logical truth is a substitution instance of a valid *schema* ([22], p. 144). This would amount to much the same as the first definition *if* a valid schema were defined as one having only true substitution instances. But that's not how he defines validity for schemata. In proving completeness, Quine relies on *interpretations* like those of what I am calling standard semantics ([22], p. 142, and Section 31, p. 176). So he defines a valid schema as

one that comes out true on all *interpretations*, which for Quine means on all assignments of *extensions* to predicate *schemata* ([22], p. 129).

This is a pointless deviation from what I am calling standard semantics. (By calling standard semantics in my sense "standard", I don't mean to imply that it came *first*.) If we are going to make use of assignments of extensions, we might as well let them be assignments to actual predicate constants, as in standard semantics, and skip the detour through schemata. The detour through schemata might have some worth if we understood what, intuitively, an assignment of extensions to predicate schemata is supposed to represent and understood it better than we understand what an assignment of extensions to predicate constants is supposed to represent, but I for one do not understand that better. We can think of an extension assigned to a predicate constant as an extension the predicate *might have had*. But there's no sense in which a predicate *schema* might have had an extension. This is not to say that Quine's conception of validity as essentially a matter of form is wrong. On the contrary, I am going to put forward such a conception of validity myself. My point is, rather, that if we want to think of validity as essentially a matter of form, then we should not define the validity of forms in terms of assignments of extensions to predicate schemata.

The two-step approach, though flawed, is one alternative to standard semantics. Another is the *truth-value semantics* developed by Hugues Leblanc (Leblanc and Wisdom [16], Leblanc [14] and [15]). A virtue of truth value semantics is that it is self-sufficient. That is, we do not need to resort to interpretations in the standard semantical sense to prove completeness.

To see how truth-value semantics works, let us first define a language L to illustrate it. First, the primitive vocabulary of L consists of denumerably many individual variables (henceforth, just "variables"), denumerably many individual constants (henceforth, just "constants"), up to denumerably many predicate constants (henceforth, just "predicates"), the connectives "~" and "⊃", and the parentheses "(" and ")". We call the variables, constants, and predicates *nonlogical* vocabulary. Second, we define a *formula* of L (i.e., well-formed formula) in such a way that all formulas are what are usually called closed formulas (i.e., formulas containing no free variables). Thus, if F^n is an n-place predicate and $c_1, c_2, \ldots, c_n$ are constants, then $F^n c_1 c_2 \ldots c_n$ is a formula; if A is a formula, then $\sim A$ is a formula; if A and B are formulas, then $(A \supset B)$ is a formula; for all variables x and constants c, if Ac/x (the result of substituting an occurrence of c for every occurrence of x in A) is a formula, then $(x)Ax$ is a formula; and nothing else is a formula.[4] All and only formulas of the sort $F^n c_1 c_2 \ldots c_n$ are called *atomic*.

In truth-value semantics, an *interpretation* Σ of L is simply an assignment of the truth values *truth* and *falsehood* to the atomic formulas of L.[5] A formula A of L is said to be *true on* Σ if and only if (i) A is atomic and $\Sigma(A) = $ truth; (ii) $A = \sim B$ and B is not true on Σ; (iii) $A = (B \supset C)$ and either B is not true on Σ or C is true on Σ; or (iv) $A = (x)B$ and for each constant c of L, Bc/x is true on Σ (where Bc/x is the result of substituting an occurrence of c for every occurrence of x in B). Clause (iv) identifies truth-value semantics as a species of substitutional semantics. Of course, this definition cannot be regarded as a comprehensive definition of truth in L, as a standard semantical definition of truth might be, since it does not define the truth value assigned to atomic formulas.

Still, this may be regarded as a definition of truth in L (on Σ) for compound formulas of L.

To define validity, i.e., the semantic consequence relation, in truth-value semantics, we first need to explain what it is for one set of formulas of L to be isomorphic to another.[6] Let a *rewrite function R* be a function from the individual constants of L to the individual constants of L. And let us say that a rewrite function R is *one-to-one* just in case for each constant c of L, R maps at most one constant into c. Then we say that a set S^* of formulas of L is *isomorphic* to a set S of formulas of L if and only if for some one-to-one rewrite function R, S^* is the result of substituting $R(c)$ for every occurrence of c, for every constant c, in the formulas in S. We may now define a set S of formulas of L to be *semantically consistent* (in the truth-value sense) if and only if for some set S^* isomorphic to S and some interpretation Σ, every member of S^* is true on Σ; S is *semantically inconsistent* (in the truth-value sense) otherwise. Finally, a formula A of L is said to be a *semantic consequence* (in the truth-value sense) of a set S of formulas of L just in case $S \cup \{ \sim A \}$ is semantically inconsistent.

The reason for the use of isomorphisms (as defined above) in the definition of semantic consequence, i.e., validity, is this: We need to be able to prove that the usual sort of predicate calculus is (strongly) complete with respect to truth value semantics. That is, we need to be able to show that if A is a semantic consequence of S in the truth value semantical sense, then A is *derivable* from S by means of our axioms and rules of inference. If we said nothing of isomorphic sets and said simply that a formula A is a semantic consequence of S just in case A comes out true on any interpretation on which the members of S all come out true, we would not be able to do this. In the usual predicate calculi, $(x)Fx$ is not derivable from $\{ Fa, Fb, Fc, \ldots \}$, and for good reasons. However, if we define semantic consequence in the simpler way just mentioned, then in light of clause (iv) in the definition of truth on Σ, $(x)Fx$ *will* be a semantic consequence of $\{ Fa, Fb, Fc, \ldots \}$ provided that $a, b, c, \ldots$ exhaust the constants of L. The subtler definition of semantic consequence in terms of isomorphic sets of formulas avoids this result since there will be sets of formulas isomorphic to $\{ Fa, Fb, Fc, \ldots \}$ that do not exhaust the constants of L.

To prove completeness, we proceed pretty much as in the context of standard semantics. The key step, as usual, is to show that if a set of formulas of L is proof-theoretically consistent (i.e., no contradiction can be derived from them), then it is semantically consistent (in the truth-value sense). By the truth-value semantical definition of semantic consistency, this requires us to show that if a set of formulas S is proof-theoretically consistent, then for some set S^* isomorphic to S there is an interpretation of L (in the truth-value semantical sense) such that each member of S^* is true on that interpretation. To show this, we first show that if a set S^* of formulas of L, *to which infinitely many individual constants of L are foreign*, is proof-theoretically consistent, then S^* can be extended to a maximally consistent, omega-complete set M of formulas of L. We then define an interpretation Σ that, for each atomic formula A of L, assigns truth to A if A belongs to M and assigns falsehood to A otherwise. Every member of M, and hence every member of S^*, turns out to be true on Σ. To get the desired result, it remains only to show that for every proof-theoretically consistent set S there is a proof-theoretically consistent set S^* to which infinitely many

individual constants of L are foreign and which is isomorphic to S. Given some enumeration of the individual constants of L, we can always construct S^* on the basis of S by substituting, for each n, the $2n$th constant for the nth constant. (For details see Leblanc [14].)

Truth-value semantics appears to be the semantics of choice for anyone who wishes to give a metalinguistic account of reference. It appears to avoid entanglements with the concept of reference inasmuch as it does not depend on assignments of individuals to constants and sets to predicates, and, unlike the Quinean two-step approach, it stands on its own in completeness proofs. However, I have three criticisms of truth-value semantics considered as a way of trying to avoid entanglements with reference. The point of my first two criticisms is that it is hard to see how we can avoid appealing to reference relations when we try to make sense of the various devices employed in truth-value semantics. I want to emphasize that the point of these first two criticisms is *not* that there is some formal flaw in truth-value semantics, or that reference relations are contained under some alias among the formal devices of truth-value semantics, or that truth-value semantics somehow commits us to falsehoods. However, the point of my third criticism will indeed be that truth-value semantics leads us to ascribe truth to certain patent falsehoods.

Before I begin, let me make explicit two further assumptions that will govern my discussion. The first is that among the truth-value interpretations of L there will be one special one, call it the *intended interpretation*, that assigns to each *atomic* formula of L the truth value that it actually has. My second assumption is that if we think of truth-value semantics as the right way to do semantics, then for any sentence of L, whether atomic or not, if the truth-value semantical definition of truth on an interpretation implies that that sentence is true on the intended interpretation, then that sentence must be true *simpliciter*.

My first doubt about truth-value semantics is whether there is any way to explain what an assignment of truth values is supposed to be without appealing to the concept of reference. One explanation might be that in assigning the value *truth* to an atomic formula we are supposing that individuals and sets are assigned to constants and predicates in such a way that the formula is true on that assignment. But if that's how we explain truth-value assignments, then we will be driven to invoke the reference relations we had aimed to avoid. The assignment of individuals to constants and of sets to predicates that underlies the *intended* interpretation, we will say, is the reference relation, and the assignments of individuals and sets that underlie the other truth-value assignments represent nonactual but possible reference relations. Of course, for any assignment of truth values to atomic formulas, more than one assignment of individuals to constants and sets to predicates will generate that assignment, but if we are thinking of truth value assignments in this way, then only the reference relation will *make sense* of the intended interpretation. A different way of thinking about assignments of truth values to atomic formulas might be to think of an assignment of truth to a formula as an assignment of *the world* to that formula and to think of an assignment of falsehood as an assignment of *nothing*. One problem with this is that the idea of assigning the world to a formula is not a lot easier to grasp than the idea of assigning a truth value to a formula. It might be clarified in terms of reference, but that, again, is what we want to avoid.

My second doubt about truth-value semantics *begins* with the observation that in truth-value semantics the quantifications of L may not all carry existential commitments. So first, let me say what I mean by "existential commitment". (Please allow me to define it my way.) Let us say that a universal quantification $(x)A$ carries existential commitment just in case for every instance Ac/x, if $(x)A$ is true (*simpliciter*) and Ac/x is a semantic consequence of $(x)A$, then c refers to something actual. (I allow that if $(x)A$ is false, then it carries existential commitment.) Let us say that an existential quantification $(\exists x)A$ carries existential commitment just in case for every instance Ac/x, if Ac/x is true (*simpliciter*) and $(\exists x)A$ is a semantic consequence of Ac/x, then c refers to something actual. (Notice that existential commitment in my sense is relative to a semantics—the semantics that defines the semantic consequence relation.)

As I say, in truth-value semantics the quantifications of L may not all carry existential commitments. To see this, suppose that L contains a translation of a sentence such as "Item 12 represents Shiva", which might be spoken by a museum curator. Surely such a sentence, and its translation into L, might be true even though "Shiva" and its translation do not refer to anything actual. Yet nothing in truth-value semantics prevents us from treating the translation into L of "Item 12 represents Shiva" as consisting of two constants and a two-place predicate. But truth-value semantics treats quantifiers substitutionally. Thus a truth-value semantical account of existential quantification will apparently have as a result that the translation into L of "There is something that Item 12 represents" is a semantic consequence of the translation into L of "Item 12 represents Shiva". In that case, since "Shiva" and its translation into L refer to nothing actual, existential quantifications in truth-value semantics will not all carry existential commitment.

Some substitutionalists, such as Marcus [18], think it a virtue of the substitutional interpretation of the quantifiers that it allows us to have quantifications free of existential commitment (in some sense). I do not wish to deny that this may be a virtue, and the problem I find is not that in truth-value semantics quantifications may be free of existential commitment. The problem I find is this: Whether or not freedom from existential commitment is a virtue, we will want there to be a way to *build existential commitment into* the semantics of a quantification for those occasions when we want it, and the only evident means of building existential commitment into quantifications as interpreted by truth-value semantics is to invoke reference relations. In truth-value semantics, if we want to ensure that all quantifications of a given kind carry existential commitment, the only evident way to do so is by stipulating that quantifications of that kind quantify only over *referring* expressions. In the case of universal quantifications, we may stipulate that a universal quantification is true if and only if every instantiation of it by a *referring* expression is true. In the case of existential quantifications, we may stipulate that an existential quantification is true if and only if some instantiation of it by a *referring* expression is true. If we do that, then the sentence of L that translates "There is something that Item 12 represents" will no longer fail the test for existential commitment since that sentence of L will not be a semantic consequence of the sentence of L that translates "Item 12 represents Shiva" (since the translation of "Shiva" is not a referring expression). Thus we will wind up appealing to the reference relations we had wished to avoid.

No doubt we could avoid actually using the term "referring" in the truth-value semantical definition of truth by mentioning instead a certain *set* of individual constants in terms of which the truth conditions of quantifications are defined. But then when we came to try to say what distinguished the members of that set, I don't see how we could avoid characterizing them as the individual constants that *refer* to something actual.[7]

Aside from these two doubts, there is a third and much more striking problem. This is that truth-value semantics is apparently going to ascribe truth (*simpliciter*) to certain sentences that in fact are clearly false. Suppose that L contains a translation of "Everything has a name". Certainly that sentence of English is *false*, since in fact there are things that do not have a name in any language, and likewise its translation into L is going to be false (assuming that in adding L to our stock of languages we do not give a name to everything). However, using the truth-value semantical account of universal generalizations, we can derive the conclusion that the translation of "Everything has a name" into L is *true*. By the method of isomorphisms, we've fixed things so that from the translations of "Ronald Reagan has a name", "Disneyland has a name" and so on, we cannot validly *infer* the translation of "Everything has a name". Still, we can tell by inspection, as it were, that no matter what name we substitute for the variable in the translation of "*x* has a name" the result is bound to be true.[8] So the intended interpretation (as defined above) will assign truth to each sentence of L that translates a sentence of the form ⌜*x* has a name⌝. So by the truth-value semantical definition of truth on an interpretation (clause (iv)), we may infer that the translation of "Everything has a name" is true on the intended interpretation. But if we think that truth-value semantics is the right way to do semantics, then, I am assuming, any sentence that truth-value semantics says is true on the intended interpretation must be true *simpliciter*. So if we think of truth-value semantics as the right way to do semantics, we will have to conclude that the translation of "Everything has a name" is true. Similarly, if L does not have a name for anything weighing a milligram or less, we will have to conclude that the translation of "Everything weighs more than a milligram" into L is true. These results show that truth-value semantics is not the right way to do semantics.[9] Alternatively, we might blame the language in question for containing too few names or blame any formal language that pretends to translate such sentences. Far more agreeable would be to blame our semantics.

3 Schematic semantics   We have looked at two alternatives to standard semantics. The Quinean two-step approach to validity has intuitive plausibility but does not stand on its own when it comes to proving completeness. Leblanc's truth-value semantics stands on its own but seems philosophically dubious. Fortunately, we can have the best of both by combining them! In essence, we may define a valid argument as an argument that has a valid form, as in the two-step approach, but then use something like truth-value semantics to define the validity of argument forms. I call this *schematic semantics*.

So we are going to need two languages, a language proper and a language of forms. Let L be the language whose vocabulary and grammar were defined in Section 2. Let F be another language whose vocabulary and grammar are de-

fined in exactly the same way except that the nonlogical vocabulary of F is written in boldface whereas the nonlogical vocabulary of L (let us suppose) is written in lightface. (Since I shall refer to formulas of both languages using metalinguistic schematic letters, this difference will not show up in what follows.) Let both L and F have denumerably many individual variables and denumerably many individual constants, and for each n let F and L have the same number of n-ary predicate constants. Intuitively, we are going to think of L as a language proper, which means that we think of the individual constants and predicate constants of L as genuinely meaningful *words*. Intuitively, we are going to think of F as the language of forms, which means that we are going to think of the *so-called* individual variables, individual constants, and predicate constants of F as schematic letters for which variables, constants, and predicates of L may be substituted.

Our strategy is going to be this: First we define *virtual truth* (on an assignment) for formulas of F. Then in terms of virtual truth we define *virtual consistency* and *virtual inconsistency* for sets of formulas of F. Then we define a *semantically inconsistent* set of formulas of L as a set of formulas of L that can be got by appropriate substitutions into some virtually inconsistent set of formulas of F. Finally, we say that a formula A of L is a *semantic consequence* of a set S of formulas of L just in case $S \cup \{\sim A\}$ is semantically inconsistent.

We are going to dispense with interpretations of L altogether. We are thinking of L as a language of meaningful formulas that are all already true or false, and we will not need to entertain their having different truth values from those they actually have. Instead, we will interpret F. And instead of assigning truth values (peculiar abstract entities) to atomic formulas of F, we will simply assign truths and falsehoods of L, i.e., formulas of L, to atomic formulas of F. We will assume that all formulas of L are either true or false and at least one is true and at least one is false.

Thus, let V be an assignment of formulas of L (not necessarily atomic) to *atomic formulas* of F. We define *virtual truth* for formulas of F on analogy with the way Leblanc's truth-value semantics defines truth for L. Thus: A formula A of F is virtually true on V if and only if (i) A is atomic and $V(A)$ is a true formula of L; (ii) $A = \sim B$ and B is not virtually true on V; (iii) $A = (B \supset C)$ and either B is not virtually true on V or C is virtually true on V; or (iv) $A = (x)B$ and, for each constant c of F, Bc/x is virtually true on V.

Let a *rewrite function R* (in the schematic semantical sense) be a function from the individual constants of F to the individual constants of F. A rewrite function R is one-to-one just in case for each constant of F, R maps at most one constant into it. We say that a set S^* of formulas of F is *isomorphic* to a set S of formulas of F if and only if for some one-to-one rewrite function R, S^* is the result of substituting $R(c)$ for every occurrence of c, for every constant c, in the formulas in S. We may now define a set S of formulas of F to be *virtually consistent* if and only if for some set S^* isomorphic to S and some assignment V of formulas of L to atomic formulas of F, every member of S^* is virtually true on V; S is *virtually inconsistent* otherwise.

Next we need to know what it means to say that a set of formulas of L is a substitution instance of a set of formulas of F. To this end, let a *substitution*

function **Sub** be a function from the individual variables of F to the individual variables of L, from the individual constants of F to the individual constants of L and, for each n, from the n-ary predicates of F to the n-ary predicates of L. Then we define a set S_L of formulas of L to be a *substitution instance* of a set S_F of formulas of F if and only if for some substitution function **Sub**, S_L is the result of substituting **Sub**(v) for each occurrence of v in S_F, for each nonlogical vocabulary item v of F.

We now define a set S_L of formulas of L to be *semantically inconsistent* (in the schematic semantical sense) if and only if it is a substitution instance of some virtually inconsistent set S_F of formulas of F; it is *semantically consistent* (in the schematic semantical sense) otherwise. Finally, a formula A of L is a *semantic consequence* (in the schematic semantical sense) of a set S_L of formulas of L if and only if $S_L \cup \{\sim A\}$ is semantically inconsistent (in the schematic semantical sense).

Now let us consider matters of soundness and completeness. I will assume that we have a proof theory of the usual sort (such as may be found in Leblanc [14], p. 13) for both L and F. The axiom schemata and inference rules by which we represent the proof theory for L are presumed to be identical to those by which we represent the proof theory for F. I now set out a number of truths without proof but with commentaries indicating how to get proofs where proofs are called for:

> (i) A set of formulas of F is proof-theoretically consistent if and only if it is *virtually consistent*. *Commentary*: Leblanc [14] has proved that a set of formulas of L is proof-theoretically consistent if and only if it is semantically consistent in the truth-value sense. A mere rewording of his proof gives us a proof of (i). (The completeness part of his proof was outlined in Section 2 above.) Leblanc's proof does not rely on the devices of standard semantics (for instance, it does not proceed simply by demonstrating that a set of formulas is consistent in the standard semantical sense if and only if it is consistent in the truth-value semantical sense), and so neither will the rewording.
>
> (ii) Every proof-theoretically inconsistent set S_L of formulas of L is a substitution instance of some proof-theoretically inconsistent set S_F of formulas of F. *Commentary*: Let S_F be a set from which S_L can be generated by a substitution function that assigns each nonlogical vocabulary item of F to a distinct vocabulary item of L, i.e., let the mapping be one-to-one. Using the inverse of this substitution function, we can construct a derivation of a contradiction from S_F on the model of the derivation of a contradiction from S_L.
>
> (iii) Every substitution instance of a virtually inconsistent set S_F of formulas of F is a semantically inconsistent (in the schematic semantical sense) set S_L of formulas of L. *Commentary*: This is true by definitions.
>
> (iv) Every semantically inconsistent (in the schematic semantical sense) set S_L of formulas of L is a substitution instance of some virtually inconsistent set S_F of formulas of F. *Commentary*: This too is true by definitions.

(v) For every set S_L of formulas of L, if S_L is a substitution instance of a proof-theoretically inconsistent set S_F of formulas of F, then S_L is proof-theoretically inconsistent. *Commentary*: Using the substitution function that generates S_L from S_F we can construct a derivation of a contradiction from S_L on the model of the derivation of a contradiction from S_F.

From (i), (ii), and (iii) it follows that every proof-theoretically inconsistent set S_L of formulas of L is semantically inconsistent (in the schematic semantical sense). Thus we can be sure that the usual sort of proof theory is strongly sound with respect to schematic semantics. From (i), (iv), and (v) it follows that if a set S_L of formulas of L is semantically inconsistent (in the schematic semantical sense) then it is proof-theoretically inconsistent. Thus we can be sure that the usual sort of proof theory is strongly complete with respect to schematic semantics. Since at no point in the proof of this do we appeal to interpretations in the standard semantical sense (or in any sense but the schematic semantical sense), schematic semantics is self-sufficient in the way Quine's two-step approach turned out not to be.

Now let us see how schematic semantics avoids the philosophical problems that attended Leblanc's truth-value semantics. My first doubt about truth-value semantics was whether we could explain what assignments of truth values are supposed to be without appealing to reference relations. In the context of schematic semantics no such doubt arises since we do not assign truth values. If a formula of L is true then we never have to entertain its being false, and if a formula of L is false then we never have to entertain its being true. In entertaining various interpretations of the atomic formulas of F, we do entertain a formula's having various *virtual* truth values. But there should be no temptation to explain this in terms of reference since the virtual truth value of a formula of F is just a matter of how we map the atomic formulas of F into formulas of L. No mapping into the nonlinguistic world is called for. Someone might think that we will need to appeal to reference relations in explaining what it means to say that a true formula of L *is* true, but that isn't obvious. In the context of schematic semantics we are not driven to explain truth *simpliciter* in terms of reference, as we are in standard semantics and possibly also in truth-value semantics. Since we do not need to consider alternatives to truth *simpliciter*, we do not need to invoke the reference relation as that which varies between the alternatives.

My second doubt about truth-value semantics was whether in the context of truth-value semantics we could build existential commitment into the quantifiers without leaning on reference relations. In the context of schematic semantics, however, there is an alternative means of building existential commitment into the quantifiers. In the context of schematic semantics, existential commitment may be secured by refinements in the notion of *form*. To build existential commitment into universal quantifications, we may conceive of form in such a way that an inference from $(x)A$ to Ac/x will not have a virtually valid form if for some constant n, An/x is true even though n does not refer to anything actual. To build existential commitment into existential quantifications, we may conceive of form in such a way that an inference from Ac/x to $(\exists x)A$ will not have a vir-

tually valid form if for any constant n, An/x might be true even though n does not refer to anything actual.

The way to realize formally these alterations in our conception of form would be to refine our definition of a *substitution function*, which, recall, was a function from vocabulary in F, the language of forms, to vocabulary in L, the language proper, and was used in defining substitution instances of sets of formulas in F. For instance, the way to deny that the translation of "There is something that Item 12 represents" follows from the translation of "Item 12 represents Shiva" will be to define the substitution function in such a way that the form of the latter is representable only as a one-place predication in the language of forms. And in general, we will define substitution functions in such a way that for any n-adic predicate P, if P may be truly instantiated in one of its places by a nonreferring constant and a constant c occupies that place, then (whether or not c refers) a substitution function will map an at most $(n - 1)$-adic predicate (or sentence) of F into the *sum* of P and c. (If a variable, not a constant, occupies that place, then the treatment is normal.) Of course, this will not mean that "represents" in "Item 12 represents Shiva" is *really* only a one-place predicate. We might like to say that "logically" it is only (part of) a one-place predicate. But grammatically, i.e., for purposes of defining well-formed formulas, it remains a two-place predicate.

Whether this strategy for avoiding reference in building in existential commitment will work depends on whether we can give a general characterization of the predicates requiring special treatment and can do so without invoking reference. What we have to avoid is precisely the characterization I used in the previous two paragraphs in introducing the strategy. There I characterized the problematic predicates as the predicates that may be truly instantiated by *nonreferring terms*. Here I can only indicate the direction in which we ought to look for our alternative characterization that avoids mentioning reference: The predicates requiring special treatment are those that in some way introduce a meaning or a mental content. For instance, "represents" in "Item 12 represents Shiva" serves to introduce the meaning of Item 12. Or for instance, "fears" in "Tammy Bakker fears the devil" introduces the content of Tammy Bakker's fear. I acknowledge that in thus using intentional locutions in the service of a theory of reference I am inverting the usual order of explanation, which is to identify intentional locutions in terms of reference. For instance, one of the marks of intentionality is supposed to be that truth is not preserved under substitutions of co-referring terms.[10]

Finally, schematic semantics does not commit us to the truth of patent falsehoods in the way truth-value semantics does. Schematic semantics does not have the result that the translation of "Everything has a name" is true. Nor does it turn out that the translation of "Everything weighs more than a milligram" is true if nothing that weighs a milligram or less has a name in the language. The reason is that schematic semantics does not define *truth*, or even *truth on an interpretation*, at all! It does not even define truth for compound formulas given the truth values of atomic formulas. Schematic semantics defines only *virtual truth* (on an interpretation) for the language of forms and so does not contradict our assumption that "Everything has a name" and "Everything weighs more than a

milligram" are false. We might still have a problem if schematic semantics validated an argument having one of these sentences as its conclusion and exclusively truths as premises, but by the method of isomorphisms we have fixed things so that this won't happen.

Whether this failure of schematic semantics to define truth on an interpretation is a problem depends on what the function of formal semantics should be. The one objective that is not too lofty is to have a means of demonstrating underivability. We have a certain deductive apparatus consisting of inference rules and possibly axioms, and we find that no matter how hard we try we cannot derive a certain conclusion from certain premises by means of this apparatus. We want a way of showing ourselves that the problem is not just our own stupidity but that the derivation is impossible. This involves two things. First, there has to be a certain operation (such as finding an interpretation of a certain sort) such that if we can perform it, then we may conclude that the derivation cannot be done. Second, we have to have a reason to try to perform the operation. That is, we have to be sure that if the derivation is indeed impossible, then some operation of this sort will show that. (This does not say that there has to be an algorithm for deciding whether there is a derivation, for there may be no algorithm for deciding whether the operation can be performed.)

For this purpose, schematic semantics will suffice. By the soundness theorem, we will know that an argument is underivable if we can show that it is invalid (in the schematic semantical sense). And if an argument is invalid, we can show that it is by showing that a most detailed form of the argument (i.e., a form from which the argument may be generated by a substitution function that is one-to-one) is invalid; and that we can do by finding a suitable assignment of truths and falsehoods of L to the atomic formulas of F. Moreover, we do have a reason to try to demonstrate the invalidity of arguments we cannot derive since, by the completeness theorem, we know that if an argument is underivable then it is invalid.

Schematic semantics avoids defining truth by defining *virtual* truth (on an interpretation) instead. What *is* virtual truth? One answer, of course, is that it's just what the definition says. But we would like to be able to regard that definition as a theoretical articulation of something we have some pretheoretical grasp on. Offhand, the definition of virtual truth may appear to tell us how the truth value of a formula having a given form depends on the truth values of its components, which in the case of a quantification are its instances. But that isn't right, for if it were, then a universal quantification would have to be true if all of its instances were, and that, as we have seen, is not always so.

A better view of virtual truth is to regard the definition of virtual truth as an explication of what we might call *semantical truth*. Semantical truth is truth-for-purposes-of-testing-validity, or more generally, truth-for-purposes-of-formal-semantics. Thus, the definition of virtual truth tells us how the *semantical* truth value of a formula having a given form depends on the semantical truth values of its components. For purposes of testing validity and other such semantical properties, it *is* permissible to regard a universal generalization as true if all of its instances are true — provided enough constants are foreign to the argument or set of formulas in question. That's precisely what Leblanc has shown. To dem-

onstrate that semantical truth is the pretheoretical concept we were looking for, one would want to show that we do draw at least an implicit distinction between truth and semantical truth in ordinary parlance.

4 Truth conditions

4 Truth conditions   The fact remains that we do know certain truths of the form:

<div align="center">"s" is true in L if and only if p.</div>

From this schema we may get one of those truths we know by substituting for "L" the name of some language, for the letter "s" a sentence of that language, and for the letter "p" a sentence of our own language that translates the sentence we substitute for "s". Call such truths *truth-sentences*.

Such truth-sentences (or something close) can be derived directly from a standard semantical definition of truth-in-L on an interpretation together with a specification of the intended interpretation of L. By contrast schematic semantics contributes nothing toward the derivation of truth-sentences. Thus standard semantics seems to offer an account of truth-sentences whereas schematic semantics offers none. If we reject the standard semantical account on the grounds that we do not understand the relation of reference with which it is entangled, then we owe an alternative account.

Here is my alternative account. To derive a truth-sentence for a given sentence of L, we need three things over and above the standard proof-theoretic apparatus. First, we need a translation of that sentence into a sentence of our own language. Let us assume we have that. Thus we will have a premise of the form:

<div align="center">"s" (a sentence of L) and "p" (a sentence of our language)
are translations of one another.</div>

Second, we need the following axiom schema:

<div align="center">If "s" (a sentence of L) and "p" (a sentence of our language)
are translations of one another, then "s" is true in L
if and only if "p" is true in our language.</div>

Third, we need the following two inference rules:

Quotation

$$\frac{p}{\text{"}p\text{" is true in our language.}}$$

Disquotation

$$\frac{\text{"}p\text{" is true in our language}}{p}$$

(These are rules, stated in the metalanguage, licensing inferences in the object language from and to sentences of the object language having the form ⌜"p" is true in our language⌝. The object language, in this case, is not L but the language in which the truth-sentences are expressed.) From the premise and one of

our axioms, we may draw a conclusion of the form ⌜"s" is true in L if and only if "p" is true in our language⌝. From this in turn we may derive a truth-sentence by means of Quotation and Disquotation. Let us call this manner of deriving truth-sentences, considered as a way of accounting for the truth of truth-sentences, the *translation-theoretic* account. (Notice that schematic semantics is not involved.) In the rest of this section I shall entertain a number of doubts about this account.

Friends of standard semantics may object that the translation-theoretic account cannot compete with the standard semantical account since the core of a translation is a standard semantical definition of truth for the language in question. According to Davidson [3], for instance, the primary objective of a deliberate, scientifically conducted translation would be to construct a standard semantical definition of truth for the language. Above all, what the translator wants to know, according to Davidson, are the truth conditions of native utterances.

I don't think the Davidsonian view of translation would seem so attractive if we were not antecedently committed to standard semantics. If we could free ourselves of that prejudice, I think a very different strategy would look more promising. The translator's primary objective is communication with speakers of the foreign language. A correct translation of the foreign language into our own language is one that best enables us to communicate with speakers of the foreign language. This may lead back to the idea that translation rests on a definition of truth if communication is conceived as a matter of discovering what is the case about the world on the basis of the speaker's words. But communication need not be so conceived. Instead, communication may be conceived as, broadly speaking, a matter of getting along via verbal activity. In large part, this "getting along" amounts to being able to predict what others will say and how they will respond to one another's words—or at least to narrow the range of possibilities. Such a capacity to get along, in some degree, with speakers of one's own native language may be taken for granted where the question is only how to translate and not how language is possible at all. The objective in translation is to find a mapping between sentences of one's own language and sentences of the foreign language such that substitution of sentences for sentences in accordance with that mapping converts one's capacity to get along in one's own language into a capacity to get along in the foreign language.

On this account, translation no more rests on an acquaintance with truth conditions than does one's capacity to get along in one's own language. This answers the Davidsonian objection since one's capacity to get along in one's own language does not rest on an acquaintance with truth conditions at all. To be acquainted with truth conditions, I am assuming, is to represent them in some kind of language (whether it be a public language or some kind of mentalese). And to have a capacity for language is to have a capacity for at least one language that one understands without having to represent the truth conditions of its sentences in some other language. (Otherwise, to understand any language, one would have to understand an infinite hierarchy of languages.) But a representation in a language L of the truth conditions of sentences in that same language L, I assume, can add nothing to one's understanding of L. So to have a capacity for language, one will have to have a capacity for some language such that one's capacity for

that language does not depend on an acquaintance with the truth conditions of its sentences.

Still, there are bound to be doubts about the rules Quotation and Disquotation. Suspicions are rightly aroused by the fact that in the context of standard semantics an account of validity that validates these rules so easily generates paradox. Let "τ" be our symbolic translation of "is true in our language". Let "*...*" be our symbolic equivalent to quotation marks. Thus, what we want is a sense in which the inference from A to $\ulcorner \tau^*A^* \urcorner$ (an inference conforming to the rule of Quotation) is valid and the inference from $\ulcorner \tau^*A^* \urcorner$ to A (an inference conforming to the rule of Disquotation) is valid. Suppose we define a standard semantical interpretation $\langle D, I_q \rangle$ of our augmented language to be *quotational* if and only if the following condition holds: $\ulcorner \tau^*A^* \urcorner$ is true on $\langle D, I_q \rangle$ if and only if A is true on $\langle D, I_q \rangle$. Then both the inference from A to $\ulcorner \tau^*A^* \urcorner$ and the inference from $\ulcorner \tau^*A^* \urcorner$ to A will be valid in the sense that, for every *quotational* interpretation, if the premise is true on that interpretation then the conclusion is true on that interpretation as well. Suppose, however, that for some constant c, and some quotational interpretation I_q, $I_q(c) = \ulcorner \sim \tau c \urcorner$ (the liar) $= I_q(\ulcorner^* \sim \tau c^* \urcorner)$. In that case $\ulcorner \tau^* \sim \tau c^* \urcorner$ and $\ulcorner \tau c \urcorner$ will have the same truth value on I_q, and so, by Quotation and Disquotation, $\ulcorner \sim \tau c \urcorner$ and $\ulcorner \tau c \urcorner$ will have the same truth value on I_q as well.

However, in the context of schematic semantics paradox cannot so easily arise (although, as I'll explain, it can arise). To see why not, let us first define a schematic semantical sense in which inferences conforming to the rules of Quotation and Disquotation are valid. This is fairly trivial. The first step is to augment the languages F and L with the necessary new vocabulary and to modify the sense in which a set of formulas in L may be a substitution instance of a set of formulas in F. The second step is to define a special kind of assignment such that the rules of Quotation and Disquotation, as formulated in the augmented language F, are valid in the sense that whenever the premises are virtually true on such an assignment, the conclusion is virtually true on that assignment as well. Inasmuch as we thus restrict ourselves to a special kind of assignment, this strategy for defining a sense in which Quotation and Disquotation are valid is analogous to the usual strategy for defining a sense in which the laws of identity are valid (namely by restricting ourselves to "normal" interpretations, i.e., those that assign the identity relation to "=").

Again, let "τ" be our symbolic abbreviation for "is true in our language", and let "*...*" be our symbolic equivalent to quotation marks. Again, what we want is a sense in which the inference from A to $\ulcorner \tau^*A^* \urcorner$ (Quotation) is valid and the inference from $\ulcorner \tau^*A^* \urcorner$ to A (Disquotation) is valid. The first major step is to add the new symbols, "τ" and "*...*", to both F and L, and to redefine *substitution instance*. Thus, we treat "τ" as a one-place predicate in both F and L, and treat "*...*" as in a class by itself in both F and L. Further, we reformulate the definitions of formula and individual constant in such a way that $\ulcorner^*A^* \urcorner$ is an individual constant if and only if A is a formula. (Of course, all the old individual constants remain.) Finally, we define a set S_L to be a *substitution instance* of S_F as we did before except that we stipulate that for the predicate "τ" in F we may substitute only the predicate "τ" in L and that if B of L is the result of the substitutions we make in A of F, then $\ulcorner^*B^* \urcorner$ must be the constant we substitute for $\ulcorner^*A^* \urcorner$.

The second major step is to define a special kind of assignment. Thus, let us say that an assignment V of formulas of L to atomic formulas of F is *quotational* if and only if for all formulas A of F, $V(\ulcorner \tau^*A^* \urcorner)$ is a true formula of L if and only if A is *virtually true* on V. (The definition of virtual truth remains the same as before.) Now we may define a set S of formulas of F to be *virtually quotation-consistent* if and only if for some set S^* isomorphic to S and some quotational assignment V of formulas of L to *atomic* formulas of F, every member of S^* is virtually true on V; S is *virtually quotation-inconsistent* otherwise. Now turn to the language L. Here we define a set S_L of formulas of L to be *quotation-inconsistent* if and only if it is a substitution instance (in our new sense) of some virtually quotation-inconsistent set S_F of formulas of F; it is *quotation-consistent* otherwise. Next, define a formula A of L to be a *quotation-consequence* of a set S_L of formulas of L if and only if $S_L \cup \{ \sim A \}$ is quotation-inconsistent. Clearly, $\ulcorner \tau^*A^* \urcorner$ is a quotation-conseqence of A and A is a quotation-consequence of $\ulcorner *A^* \urcorner$. Thus, we have defined a sense in which Quotation and Disquotation are valid rules of inference.[11]

Paradox cannot be generated from this schematic semantical account of the validity of inferences conforming to Quotation and Disquotation in the simple way paradox was generated from the proposed standard semantical account of Quotation and Disquotation. The reason is that schematic semantics does not involve making assignments to constants. Unfortunately, Quotation and Disquotation do present further problems. If inferences conforming to Quotation and Disquotation are unrestrictedly valid and these rules are combined with a standard (syntactic) theory of identity, then we are still liable to produce the paradox of the liar. But for two reasons I maintain that the translation-theoretic account of truth-sentences is viable even so. First, I believe it is evident that Quotation and Disquotation are in some sense valid (though perhaps not unrestrictedly). Second, the liar paradox cannot force us to adopt a standard semantics, because even from the standpoint of standard semantics the liar is a problem. Surely we cannot be satisfied with the usual expedient of postulating an infinite hierarchy of metalanguages.

5 Conclusion   We have seen that in adopting a standard semantics we are driven to seek an analysis of reference instead of merely accounting for the correct use of "refers". Thus if we are dissatisfied with all attempts to analyze the concept of reference we will be driven to seek a semantics that does not call for an analysis of reference. Our search for such a semantics led to what I call *schematic semantics*, which, in essence, defines validity in terms of logical form and uses a substitutional interpretation of the quantifiers in defining the validity of logical forms. An apparent problem with schematic semantics is that it does not yield truth conditions. We have now seen how truth conditions might be explicable without the aid of a standard semantics.

<div align="center">NOTES</div>

1. These inference rules cannot be turned into analyses. For instance, we must not put forward as an analysis of the reference of singular terms that "t" refers to a if and only if for some b, $a = b$ and "b" is a translation of "t". Such an analysis would in-

volve a simultaneous quantification into and out of quotation marks. Unless certain precautions are taken, the practice of simultaneously quantifying into and out of quotation marks is going to yield nonsense. Consider, for instance, the fallacy of inferring "Something is hungry and, with quotation marks around it, is composed of four letters" from "John is hungry and 'John' is composed of four letters". The nonsense can be avoided by treating quotation as a function from objects to their canonical names. Then in place of the fallacy just cited we have an unobjectionable inference to "Something whose name is composed of four letters is hungry". But an explication of this quotation function would be in effect a theory of reference, which is what we are supposed to be providing. To avoid the nonsense it is not sufficient, as Hill ([10], p. 201) seems to think it is, simply to declare that the quantification is to be taken as substitutional. The substitutional interpretation must not be treated as a way of making sense of what is patently nonsense when expressed in ordinary English.

2. Of course, it is not only the reference relation that needs to be clarified if we are to think of truth in a natural language as like truth-on-an-interpretation for a formal language. Further, we need to get straight about the *range* of the reference relation. For instance, if we want to preserve standard set theory, then the set of all individuals in reality must not itself be a member of the set of all individuals in reality. And of course we need to deal with all those devices of natural language, such as adverbs, that are difficult to deal with in a formal way.

3. For more on this condition and the question of its sufficiency, see the debate between Putnam [20], on the one side, and Devitt [6] (Section 11.4) and Lewis [17], on the other.

4. Here and wherever possible I dispense with corner quotes.

5. I am simplifying Leblanc's presentation in [14] by excluding partial interpretations, i.e., interpretations that make assignments to only some of the atomic formulas of L.

6. The definition of semantic consequence that I am here embarking upon is the definition Leblanc employs most extensively in [14], and it is the most congenial to later developments in the present paper. But there are other possibilities in line with the basic idea of truth-value semantics. See [14], Section 2.4, and [15] for alternatives. Yet another approach, which differs somewhat more drastically from the one taken here, may be found in Hugly and Sayward [11].

7. In other words, we could get around the problem by stipulating that a universal quantification $(x)A$ is true on a truth-value semantical assignment Σ if and only if for every constant c belonging to a set of constants S, Ac/x is true on Σ. (Accordingly, an existential quantification $(\exists x)A$ would be true on Σ if and only if for some constant c belonging to S, Ac/x is true on Σ. The truth of "Item 12 represents Shiva" will then fail to guarantee the truth of "There is something that Item 12 represents" inasmuch as "Shiva" may not belong to S.) This is precisely the move that Leblanc makes in [14], p. 135, in the course of defining a truth-value semantics for presupposition-free logic. To verify that the temptation to explain this device in terms of reference is real, I would like to point out that Leblanc then writes: "Readers who care to may thus think of [S]. . . as comprising all and only those individual parameters . . . which, as regards [Σ], stand for *designating terms*" ([14], p. 136). Of course, this implies no criticism of Leblanc, since his aims may not be mine.

8. Maybe this isn't as obvious as I here take it to be. Since it isn't obvious that Pete Rose has a name in, say, Urdu, why should it be obvious that Pete Rose has a name

in *any* language? Apparently, if it is obvious that Pete Rose has a name in *some language*, then that is only because it is obvious that Pete Rose has a name in *this* language, the one we are using now. Thus the possibility arises that the problem that "Everything has a name" poses to truth-value semantics will disappear when we have somehow resolved the paradoxes of self-reference. But I think the problem can be posed without raising the specter of self-reference. Without making any kind of self-reference I think we can accept that every instance of the schema ⌜"x" is a name of x⌝ is true. Thus for any name x, we may derive ⌜x has a name⌝ from a corresponding instance of that schema. The derivation will be analogous to the derivation of ⌜x has a brother⌝ from ⌜y is a brother of x⌝. Incidentally, beware of assuming that "has a name" means "is assigned to some constant on every model of the theory to which this sentence belongs". It doesn't mean that.

9. These results are what Kripke [13] thinks are the ultimate flaw in the substitutional interpretation.

10. Someone might object that if the class of problematic predicates can be characterized without appealing to reference relations, then that characterization can also be used in the context of truth-value semantics to build existential commitment into the quantifiers. The idea might be that given this characterization we could simply forbid quantificational binding of the problematic places in the problematic predicates. But, on the contrary, in order to secure existential commitments we should not have to deny that a formal-language equivalent of a sentence such as "There is an x such that item 12 represents x" is well-formed. Or someone might object that in the truth-value semantical definition of truth we could revise the clauses for quantifications so that if x occupies one of the problematic places in one of the problematic predicates in A, then (∃x)A is false on Σ. But, on the contrary, we shouldn't have to suppose that existential quantifications binding problematic places in problematic predicates are always false. We just want to deny that existential quantifications that bind problematic places in problematic predicates *follow* from their instances.

11. A metalinguistic account of reference such as I proposed in Section 1 will naturally go together with an account of truth that eschews analysis of truth in favor of merely explaining the correct use of "true". Surely an integral part of such an account of truth would be that Quotation and Disquotation be in some sense valid. Thus someone might object that I must not speak of truth at all in accounting for the validity of Quotation and Disquotation, lest the account of truth to which my account of reference is tied be circular. (Putnam makes such an objection to the redundancy theory of truth in the introduction to [21].) But in fact we may define the relevant sort of validity in terms of truth without introducing circularity into our accounts of truth and reference. That's because the accounts of truth and reference, as I envision them, will actually *list* the rules governing the use of "true" and "refers" and will not need to *mention* at all the kind of validity they have.

REFERENCES

[1] Bergman, M., J. Moor, and J. Nelson, *The Logic Book*, Random House, New York, 1980.

[2] Brandom, R., "Reference explained away," *The Journal of Philosophy*, vol. 81 (1984), pp. 469–492.

[3] Davidson, D., "Radical interpretation," *Dialectica*, vol. 27 (1973), pp. 313–328.

[4] Davidson, D., "Reality without reference," *Dialectica*, vol. 31 (1977), pp. 177–185.

[5] Devitt, M., *Designation*, Columbia University Press, New York, 1984.

[6] Devitt, M., *Realism and Truth*, Princeton University Press, Princeton, NJ, 1984.

[7] Field, H., "Tarski's theory of truth," *The Journal of Philosophy*, vol. 69 (1972), pp. 347–375.

[8] Grover, D., J. Camp Jr., and N. Belnap Jr., "A prosentential theory of truth," *Philosophical Studies*, vol. 27 (1975), pp. 73–125.

[9] Hartman, T., "The heterodox interpretation of reference talk," *Philosophical Studies*, vol. 46 (1984), pp. 145–169.

[10] Hill, C., "Rudiments of a theory of reference," *Notre Dame Journal of Formal Logic*, vol. 28 (1987), pp. 200–219.

[11] Hugly, P. and C. Sayward, "Do we need models?," *Notre Dame Journal of Formal Logic*, vol. 28 (1987), pp. 414–422.

[12] Klenk, V., *Understanding Symbolic Logic*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[13] Kripke, S., "Is there a problem about substitutional quantification?," pp. 325–419 in *Truth and Meaning*, edited by G. Evans and J. McDowell, Clarendon Press, Oxford, 1976.

[14] Leblanc, H., *Truth-Value Semantics*, North Holland , New York, 1976.

[15] Leblanc, H., "Alternatives to standard first-order semantics," pp. 189–274 in *Handbook of Philosophical Logic, Vol. I: Elements of Classical Logic*, edited by D. Gabbay and F. Guenther, Reidel, Boston, 1983.

[16] Leblanc, H. and W. Wisdom, *Deductive Logic*, Allyn and Bacon, Boston, 1972.

[17] Lewis, D., "Putnam's paradox," *Australasian Journal of Philosophy*, vol. 62 (1984), pp. 221–236.

[18] Marcus, R., "Quantification and ontology," *Noûs*, vol. 6 (1972), pp. 240–250.

[19] McDowell, J., "Physicalism and primitive denotation: Field on Tarski," *Erkenntnis*, vol. 13 (1978), pp. 131–152.

[20] Putnam, H., "Models and reality," *The Journal of Symbolic Logic*, vol. 45 (1980), pp. 464–482. (Reprinted in [21], pp. 1–25.)

[21] Putnam, H., *Realism and Reason: Philosophical Papers, Vol. 3*, Cambridge University Press, Cambridge, MA, 1986.

[22] Quine, W., *Methods of Logic*, Third Edition, Holt Rinehart and Winston, New York, 1972.

[23] Quine, W., *Philosophy of Logic*, Second Edition, Harvard University Press, Cambridge, MA, 1986.

[24] Ramsey, F., "Facts and propositions," *Proceedings of the Aristotelian Society*, Supp. vol. 7 (1927), pp. 153–170.

[25] Sellars, W., *Science and Metaphysics*, Humanities Press, New York, 1968.

Department of Philosophy
University of Cincinnati
Cincinnati, Ohio 45221