DALE GOTTLIEB AND TIMOTHY McCARTHY

SUBSTITUTIONAL QUANTIFICATION AND SET THEORY

Our concern in this paper is to defend the use of substitutional quantification in set theory as a way of avoiding ontological commitment to sets. Specifically, two objections to this procedure are addressed. (1) Charles Parsons claims that substitutional quantification (at least in set theory) is not ontologically neutral, but rather expresses a *bona fide* sense of existence.[1] We argue that he has failed to distinguish between meta-linguistic commitment to expressions on the one hand and ontological commitment to sets in the object language on the other. (2) T. S. Weston claims that a substantial interpretation of the quantifiers of Zermelo–Frankel set theory (ZF) is inconsistent with obvious theses of semantics.[2] We argue that he has artificially limited the ways in which the quantification of ZF can be rendered substitutional due to a misunderstanding of the finiteness requirements for semantics. With the limitation removed, we give an example of a substitutional interpretation of ZF which is consistent if ZF itself is. It must be noted that in this discussion we are concerned only to defend the substitutional interpretation of the quantifiers. Thus we allow ourselves use of a referential interpretation of the atomic sentences of set theory; that is, 'ϵ' will be interpreted as a relation over an ontology of sets. Of course, this interpretation of the atomic sentences subverts any overall effort at ontological neutrality as regards sets. Accordingly, we shall conclude with a description of a semantical framework within which an interpretation of the language of set theory which is not committed to sets is possible.

I

The classical case for the ontological neutrality of substitutional quantification was made by Quine. He argued that (a) since substitutional quantification is explicable in terms of truth and substitution no matter what the substitution class – even that whose sole member is the left-hand parenthesis – we must deny existential import to substitutional quantification. For we surely do not have to provide an entity for the left-hand

Journal of Philosophical Logic 8 (1979) 315–331. 0022–3611/79/0083–0315$01.70.

parenthesis to name in order to count true sentences that contain it. And (b) " . . . we [cannot] introduce any control by saying that only substitutional quantification in the substitution class of singular terms is to count as a version of existence . . . [because] the very notion of singular terms appeals implicitly to . . . objectual quantification".[3] Parsons, by contrast, claims that

> . . . the existential quantifier substitutionally interpreted has a genuine claim to express a concept of existence which has its own interest and which may offer the best explication of the sense in which 'linguistic' abstract entities – propositions, attributes, classes in the sense of extensions of predicates – may be said to exist.[4]

Concerning set theory in particular, he asserts:

> The fact that the substitution interpretation yields truth conditions for quantified sentences [of set theory] means that everything necessary for speaking of these classes as entities is present, and the request for some more absolute verification of their existence seems senseless.[5]

Thus while not suggesting that substitutional quantifiers be understood objectually, Parsons does reject their ontological neutrality. He must therefore answer Quine's argument. He does so as follows: Parsons replies to (a) by pointing out two formal features of substitutional quantification with respect to singular terms which he thinks can be used to distinguish that substitution class from other (trivial) ones. They are the fact that class "admits identity with the property of substitutivity *salva veritate*. [And] it has infinitely many members that are distinguishable by the identity relation".[6] We find this appeal unconvincing. The role of identity in a substitutionally interpreted language is quite different from its ordinary role (as Parsons himself seems to recognize in his footnote 8). As Ruth Marcus pointed out,[7] identity will be replaced by a series of syntactically defined substitution principles depending for their scope on the expressive richness of the language. We may expect such principles for predicates and operators in addition to singular terms, and there seems to be no reason why we could not have them for punctuation devices as well. This suffices to cast doubt on both features to which Parsons appeals, for any expressions for which we formulate such principles could belong to an infinite class of expressions which are distinguished by those principles.

Parsons' response to (b) is to

> concede Quine's point . . . for a certain central core class of singular terms, which we

might suppose to denote objects whose existence we do not expect to explicate by substitutional quantification. We might then make certain analogical extensions of the class of singular terms in such a way that they are related to quantifications construed as substitutions. The criterion for 'genuine reference' is given in other terms.[8]

It seems to us that this does not suffice to answer (b). Quine's original problem was that there is no purely syntactic criterion for terms which are used to denote objects. Expressions which grammar classifies as proper names or definite descriptions are often used without commitment to the existence of entities for them to name. To identify the genuinely referential uses of terms, Quine suggested that we use the unquestionably referential quantifier '$(\exists x)$' as a touchstone: a term t as used in a context $\ulcorner \ldots t \ldots \urcorner$ is referential iff $\ulcorner (\exists x)(x = t) \urcorner$ follows from $\ulcorner \ldots t \ldots \urcorner$. Notice that there is no pretense to an *explication* of existence here; rather, the quantifier functions as a reliable *symptom* of referential use. Now what is accomplished by Parsons' 'analogical extensions'? They may give us reason to classify class abstracts with the singular terms syntactically – they replace similar variables, and so on. But it is their semantical interpretation which is in question: Do they denote? Quine's test via the objectual '$(\exists x)$' gives a negative answer. To appeal to the substitutional quantifier '$(\exists x)$' is to beg the question: the newly introduced 'singular terms' were supposed to *establish* the 'concept of existence' which that quantifier expresses. And in any case, such an appeal will fail. For, according to the substitutional interpretation,

$$\exists x(\text{the museum has a statue of } x)$$

is true even if the museum's only statue is a statue of Zeus, and yet we do not accord to Zeus *any* kind of existence. For the same reason, the truth of

$$\exists \alpha(\forall x(x \in \alpha \equiv x \text{ is a horse}))$$

should not lead us to accord any sense of existence to $\{x : x \text{ is a horse}\}$.

Perhaps Parsons is relying on the presumption that quantification in the meta-language is objectual and so the truth condition for quantified sentences engenders ontological commitment to expression-types. This would explain why he limits his theory to what he calls 'linguistic' abstract entities, e.g., propositions, attributes and extensions of predicates. Hence also his worry at the end whether commitment to all abstract objects has been avoided. If this is what Parsons intends, then perhaps his explication of existence for these entities goes as follows. 'Ordinary' existence is what is required for the truth of existentially quantified sentences, when the quantifier is interpreted in the ordinary manner. *Linguistic existence* – our name for what Parsons is explicating if this interpretation of his intentions is correct – is what is required for the truth of existentially quantified sentences, when the quantifier is interpreted substitutionally *in an objectual meta-language*. But then what is the difference between the linguistic existence of classes, say, and the ordinary existence of predicates? From

'$\exists\alpha\phi\alpha$' is true

it will follow, in the meta-language, that

(#) there is a predicate 'F' such that '$\phi\{x : Fx\}$' is true,

where 'there is' expresses 'ordinary' existence. Thus from the point of view of ontological commitment, the linguistic existence of classes amounts to the ordinary existence of expressions. If so, we think it is misleading to call it a new sense of existence and announce it as a metaphysical innovation.

Furthermore, even this weakened reading of Parsons' claim is open to two objections. First, his assumption that the meta-language in which the substitutional quantification of the object language is interpreted must employ objectual quantification is questionable.[9] Second, it is incorrect to associate a sentence S with the ontological commitments of the sentence S^* of the meta-language which gives S's truth-condition. S^* may have ontological commitments due solely to the structure of the meta-language and which are entirely unconnected to the content of S. For example, S^* may refer to sequences of objects (as is typically the case for Tarski-type truth theories) and hence require the existence of sets, even if S is 'Fido barks'; but 'Fido barks' is uncommitted to sets. Likewise, the fact that (#) will be the meta-language truth condition for '$(\exists\alpha)\phi\alpha$' does not allow us to transfer the former's commitment to the existence of expression types to the latter. And in this case the mistake is even more obvious: since the meta-language is presumed not to have substitutional quantification, a homophonic assignment of truth conditions is impossible, and so we expect some distortion of content. We conclude, then, that Parsons has not refuted the ontological neutrality of substitutional quantification.

II

While objections to the ontological claims of proponents of substitutional quantification have come from many quarters, arguments purporting to demonstrate the impossibility of a substitutional construal of quantification in particular theories have been comparatively rare. An objection of the latter kind, however, has recently surfaced in a paper by T. S. Weston, who argues that if we presuppose Davidson's well-known thesis that no language can count as learnable if infinitely many of its expressions must be treated as semantically primitive, then a substitutional interpretation of quantifiers in Zermelo–Frankel set theory (ZF) is inconsistent with principles of semantics common to both the referential and substitutional theories of quantification. In this section, we shall describe and criticize the argument for this contention. That argument, while developed around the case of ZF, is based upon certain general semantic assumptions which it will have value to develop and criticize.

As usually axiomatized, ZF has the set-membership predicate as its sole extra-logical primitive. In order to make it amenable to substitutional semantics, therefore, ZF must be supplemented with singular terms. Weston's contention is that there is a preferred way of doing this given the Davidson learnability stricture, and that the systems which result, when supplemented with a few trivial conditions on truth which are unexceptionable with respect to substitutional semantics, will prove that each substitution instance of some formula is true, but that the (universal) closure of that formula is false.[10] The difficulty here is in fact somewhat more pervasive than Weston apparently realizes. In analogy to the number-theoretic case, we shall say that a substitutionally interpreted theory T is 'ω-inconsistent' if, for some formula ϕ, T proves every instance of ϕ relative to the intended substitution class, but refutes the universal closure of ϕ. The extensions of ZF which Weston identifies as favored by Davidson's condition are those which result from ZF by the addition of finitely many function symbols and finitely many individual constants, where an individual constant is understood to be a singular term having no first-order structure (relative to the first-order syntax of the language of the relevant extension). It may be shown that any such extension, irrespective of supplemental semantical axioms, is ω-inconsistent, and thus obviously not substitutionally interpretable. In addition, the ω-inconsistency accrues not only to the relevant extensions of ZF, but generally to a range of weaker systems as well. The reason, in outline, is as follows. Consider, for definiteness, a system E such that the language of E (L_E) results from L_{ZF} by addition of a single individual constant 'c' and function symbol 'f'. (It will be seen that the argument to be sketched may be generalized to E's with singular terms generated from any finite number of individual constants via application of any finite number of function symbols.) Suppose that arithmetic is relatively interpretable in E, and that E represents the following simple inductive definition:

(5) $\phi(0) = c; \quad \phi(n') = f(\phi(n)).$

That is to say, there is a formula Bxy of two free variables in L_E such that

(i) $\vdash_E B(0, c), \vdash_E B(0', f(c)), \vdash_E B(0'', f(f(c))), \ldots$

and such that

(ii) $\vdash_E \forall x \exists y (Bxy \;\&\; (z)(Bxz \rightarrow z = y))$.[11]

If E assumes the axioms of infinity and replacement, by virtue of (ii) E proves that the range of the function determined by Bxy on the integers exists; i.e., using some obvious abbreviation,

(6) $\vdash_E (\exists y)(y = \{x\text{: } (\exists z)(\text{'}z \text{ is an integer'} \;\&\; Bzx)\})$.

Now it is easily seen that our assumptions on E guarantee (e.g., via the argument of Russell's paradox) the existence of a proof in E of the sentence

(7) $\forall x \sim \forall y\, y \in x$

and so, by specification,

(8) $\vdash_E \sim \forall x(x \in \{x\text{: } (\exists z)(\text{'}z \text{ is an integer'} \;\&\; Bzx)\})$.

But by (i) and (6), if 't' represents any closed singular term of E,

(9) $\vdash_E t \in \{x\text{: } (\exists z)(\text{'}z \text{ is an integer'} \;\&\; Bzx)\}$.

(8) and (9) imply that E is ω-inconsistent. Bringing matters together, then, we see that ω-inconsistency accrues to any system which assumes the axioms of infinity and replacement, and in which inductive definitions of the most elementary kind are representable *if the singular terms of that system are obtained in the manner in which Weston supposes they must be obtained*. We now turn to the argument for this supposition.

Of the various strategies for augmenting ZF with a class of singular terms suitable for a substitutional construal of its quantifiers, Weston considers only (a) adding infinitely many individual constants to L_{ZF}, and (b) adding finitely many function symbols and finitely many individual constants to L_{ZF}. Weston asserts that any language resulting from that of ZF via strategy (a) is unlearnable, or, at the very least, that such a supplementation yields a language "much less suitable for use by human beings" than one obtained via strategy (b).[12] The principle motivating this assertion was first described by Davidson: no language can count as *learnable*, Davidson writes, if infinitely many of its expressions are semantically primitive, where a semantically primitive expression of a language **L** is one such that the rules which give the truth-conditions for the sentences in which it does not appear do not suffice to determine the truth-conditions of those sentences in which it does appear.[13] Semantic primitiveness for **L** is thus characterized as relative to a set of semantical rules which associate each sentence of **L** with a representation of its truth-conditions; for example, a recursive theory of truth in the manner of Tarski, or a translation scheme mapping **L** into a language **L***, supplemented with a semantic description of **L***. Now Weston focuses exclusively on the latter construal, in terms of translation, although it should be remembered that the finiteness constraint with respect to translation is only a part of the broader requirement of finiteness applying to a complete assignment of truth-conditions to the sentences of **L**, of which the translation of **L** into **L*** is only a part. In any case, fixing English as the interpreting language, Davidson's learnability stricture for translation assumes the form

(D) L is learnable only if there is a translation function mapping the sentences of L onto sentences of English which treats only finitely many expressions of L as semantically primitive.

where a *translation* treats an expression as semantically primitive if and only if the rules of translation determining the images of the sentences in which it does not appear do not suffice to determine the images of those sentences in which it does.[14] Thus understood, it follows that the present constraint is satisfied by any translation scheme incorporating only finitely many effective rules, i.e., by any translation which is recursively specifiable. But from the condition that a translation of **L**, if satisfactory, must be recursive – the requirement that **L** possess some finitely representable structure underlying translation – it does not follow that **L** may contain only finitely many expressions which are *primitive with respect to first-order syntax in L*.

To see this, consider, as **L**, the language of Quine's theory of virtual classes as supplemented with substitutional quantifiers taking the class abstracts as substitutes. We shall assume that this theory is formulated for a referentially interpreted language, $\mathbf{L}_0$, which admits recursive translation into English; thus, if 'F' is a predicate of $\mathbf{L}_0$, let 'F^*' represent its English translation. The formulas of **L** are now generated by first-order constructions from the basic formulas, viz., those of the form

$$(\alpha) \qquad t_i \in \{x\colon Fx\},$$

where 't_i' represents a singular term or variable of $\mathbf{L}_0$ and 'F' a predicate of one free variable of $\mathbf{L}_0$, or of the form

$$(\beta) \qquad t_i \in s_i$$

where 't_i' is as above and 'v' is a substitutional variable replaced by the class abstracts '$\{x\colon Fx\}$'. Notice that the infinitely many abstracts of this form have no first-order structure *in* **L**, and must therefore be counted as individual constants with respect to **L**. Nevertheless, even though containing infinitely many individual constants, **L** is clearly susceptible of effective translation into (regimented) English, thus: the variables 't_i' are translated via corresponding variables 'u_i' ranging over the ontology of $\mathbf{L}_0$, and the variables 's_i' via corresponding variables 'v_i' ranging over sets definable in English by (translations of) one-free variable predicates of $\mathbf{L}_0$. (It is assumed that the ontology of $\mathbf{L}_0$ forms a set, although this is not essential.) Formulas of the form (α) are now interpreted by the corresponding formulas

$$(\alpha^*) \qquad \exists y(\forall x(x \in y \longleftrightarrow F^*x) \;\&\; u_i \in y)$$

and formulas of form (β) by

$$(\beta^*) \qquad u_i \in v_i.$$

The translation is extended inductively to all formulas of **L** by interpreting the connectives and quantifiers of **L** by means of corresponding connectives and quantifiers of English; in particular, the substitutional quantifiers of **L** are construed via quantifiers of English ranging over the sets definable by one-free-variable formulas of $\mathbf{L}_0$. The present example shows that (D) is not sufficient by itself to motivate the claim that translations defined on languages with finitely many first-order primitives possess 'learnability' properties which translations defined on languages with an infinite number of such primitives lack; for the translation just (recursively) specified does not treat infinitely many expressions of **L** as semantically primitive, even though **L** harbors an infinite number of singular terms having no first-order structure (*in* **L**). The general point here illustrated is that, although the learnability of a translation by a finite being does presumably require that it be finitely representable, such a representation may require a grip on more structure in the object language than its first-order structure.

However, Weston appears, at one point, to recognize that a translation function may be finitely specified over a language with infinitely many individual constants, and so satisfy (D), but suggests that languages with infinitely many constants may be inadmissible in other respects:

> In effect, the infinitely many constants of such a language are treated notationally as if they were semantically primitive, and it is only the translation which shows that they are not. Following [this strategy] in arithmetic, that is replacing '0' and ''' by an infinite list of individual constants, has obvious disastrous consequences. Since the successor function would no longer be representable, the axioms could no longer be stated in the usual way. In following strategy (4) [our (2), viz., using finitely many function symbols and constants] – either on arithmetic or ZF – we are able to represent the meaning relations which makes it possible to translate them into English.[15]

It is not obvious what Weston intends by 'meaning relation' in this context, but it seems plausible to suppose that what is meant is simply any semantically relevant structure of the object language. The response would then be that, even if, by strategy (2), we can render such meaning relations finitely 'representable' in the object language, it does not follow that this strategy is necessary to ensure such representability. The point can be illustrated by considering Weston's own example of elementary arithmetic (Z). One familiar formulation of Z employs addition and multiplication predicates $\Sigma(x, y, z)$ ("z is the sum of x and y"), $\pi(x, y, z)$ ("z is the product of x and y"), an individual constant '0' and the successor functor 'Sc'.

Weston asks us to imagine a substitutional arithmetic Z^* using the addition and multiplication predicates and an infinite set of individual constants $\{a_0, a_1, \ldots\}$ replacing {'0', 'Sc(0)', . . . }. It is not the case, as Weston suggests in the above passage, that the usual axiomatic formulation is barred because the successor functor is not representable. If a_1 is the constant replacing 'Sc(0)', one may construe '$y = \mathrm{Sc}(x)$' by $\Sigma(y, x, a_1)$. In fact, we may use this construal of the successor relation to specify a translation τ mapping L_Z into L_{Z^*} relative to which Z is relatively interpretable in Z^*,[16] where Z^* is axiomatized via the τ-translates of the axioms of Z.

First, note that any singular term of L_Z is generated from variables or '0' by iterated applications of the successor functor. By the *rank* of a singular term t of L_Z we shall understand the number of occurrences of the successor functor in t. We now define a mapping of singular terms t of L_Z onto formulas $\tau_t(x)$ of L_{Z^*} inductively on the rank of t. If t is of rank 0, then $\tau_t(x)$ is $\ulcorner x = a_0 \urcorner$, if t is '0', and is $\ulcorner x = t \urcorner$ if t is a variable. If t has rank $n + 1$, then $t = \mathrm{Sc}(t')$ for some t' of rank n. $\tau_t(x)$ is then the formula

$$\forall y(\tau_{t'}(y) \rightarrow \Sigma(x, y, a_1)).$$

If t is closed, then, the formula $\tau_t(x)$ inductively represents the denotation condition of t and if t is open, $\tau_t(x)$ inductively represents its assignment condition. The definition of ϕ^τ is now given as follows. If ϕ is $B(t_1, \ldots, t_n)$ where B is a primitive predicate of L_Z, ϕ^τ is the formula

$$\forall x_1 \ldots \forall x_n \left(\bigwedge_{j<n} \tau_{t_j}(x_j) \rightarrow B(x_1, \ldots, x_n) \right).$$

If ϕ is $\ulcorner \psi \vee \theta \urcorner$, ϕ^τ is $\ulcorner \psi^\tau \vee \theta^\tau \urcorner$. If ϕ is $\ulcorner - \psi \urcorner$, ϕ^τ is $\ulcorner - \phi^\tau \urcorner$, etc.; if ϕ is $\ulcorner \exists v\ \psi v \urcorner$, ϕ^τ is $\ulcorner \exists v\ \psi^\tau \urcorner$, and if ϕ is $\ulcorner \forall v\ \psi v \urcorner$, ϕ^τ is $\ulcorner \forall v\ \psi^\tau \urcorner$.

The axioms of Z^* are now simply the τ-translates of the axioms of Z. The relative interpretability result follows easily by induction on the length of proofs in Z.

Finally, a rather subtle response to this rejoinder must be examined. It is true that Z is relatively interpretable in Z^* by means of the above translation. That is, there is a translation of L_Z into L_{Z^*} that preserves first-order deducibility. However, there are still semantic relations that escape representation in Z^*. There are true atomic sentences of the forms $\Sigma(a_i, a_j, a_k)$ and $\Pi(a_i, a_j, a_k)$ that are not theorems of Z^*. However, since any true atomic sentence of L_Z is a theorem of Z, the τ-translate of any true atomic sentence of Z is a theorem of Z^*. We see, then, that although any atomic sentence of Z^* is equivalent to the τ-translate of a certain atomic sentence of Z, this equivalence may not be representable in Z^*, i.e., it will typically not be representable by a first-order derivation in Z^*.

Now the objector is surely correct in maintaining that it is a clear constraint on our overall semantic description of L_{Z^*} that it account for the existence of such relations. The mistake, we suggest, comes in identifying the capacity of a semantics for L_{Z^*} to provide for an entailment relation in L_{Z^*} with the capacity to explain the existence of that relation by appeal to the semantic analysis of its first-order structure (i.e., the substitutional recursion); for this is the effect of requiring that the relation be representable by means of a first-order derivation in L_{Z^*}. However, the substitutional truth-clauses provide only a part of the semantic description of L_{Z^*}. To complete matters, we must specify an assignment of truth-conditions to the atomic sentences. What is required is that the substitutional recursion in conjunction with such an assignment impose enough semantic structure to enable the logical relations among the sentences of L_{Z^*} to be mapped. (If they do not, then, there is no reason *a priori* to level the complaint against the interpretation of the quantifiers rather than the semantics for the atomic sentences.)

As an illustration, consider an assignment of truth-conditions to the atomic sentences of L_{Z^*} based upon the Zermelo interpretation of number theory in ZF. The addition and multiplication predicates may be associated, in one of the usual ways, with formulas $A(x, y, z)$ and $B(x, y, z)$ of L_{ZF} that respectively give their satisfaction conditions. The individual constants $\{a_0, a_1, \ldots\}$ are effectively mapped onto a set $\{\psi_0(x), \psi_1(x), \ldots\}$ of one-free-variable formulas of L_{ZF} that give their denotation conditions. More specifically, if $\psi_1(x)$ gives the denotation condition of a_1 [17] and for any n, if $\psi_n(x)$ gives the denotation condition of a_n, we may put

$$\text{(10)} \qquad \forall x(\psi_{n+1}(x) \longleftrightarrow \forall y \forall z[\psi_n(y) \to (\psi_1(z) \to \Sigma(x, y, z))]).$$

An atomic sentence $\Phi(a_{i_1} \ldots a_{i_n})$ of L_{Z^*} is therefore accorded

$$\forall x_1 \ldots \forall x_n\left(\bigwedge_{j<n} \psi_{i_j}(x_j) \to \Phi^*(x_1, \ldots, x_n)\right)$$

as its ZF-translation, where Φ^* is the ZF-translation of Φ, a primitive predicate of L_{Z^*}.

Now let $\Phi(x_1, \ldots, x_k)$ be any formula of L_Z. Our problem is to derive the equivalences

$$\text{(11)} \qquad [\Phi(\bar{n}_1, \ldots, \bar{n}_k)]^\tau \longleftrightarrow \Phi^\tau(a_n, \ldots, a_{n_k})$$

from the substitutional truth theory for L_{Z^*} in conjunction with the interpretation of the atomic sentences of L_{Z^*} in L_{ZF} that we have specified

A. For any n, the formula

$$(12) \qquad \forall x(\Sigma(x, a_n, a_1) \longleftrightarrow x = a_{n+1})$$

is true in L_{Z^*}.

Proof. From the definition of $\{\psi_i\}$, it follows by induction on n that

$$\forall x(A(x, a_n^*, a_1^*) \longleftrightarrow \psi_{n+1}(x)),$$

where for any i, a_i^* is the denotation of a_i in the world of sets. (12) now follows by the equivalences

$$\forall i \forall j \forall k(A(a_i^*, a_j^*, a_k^*) \longleftrightarrow \Sigma(a_i, a_j, a_k) \text{ is true in } L_{Z^*})$$

and

$$\forall n \forall x(x = a_{n+1}^* \longleftrightarrow \psi_{n+1}(x)).$$

B. For any n, the formula

$$(13) \qquad \forall x(\tau_{\bar{n}}(x) \longleftrightarrow x = a_n)$$

is true in L_{Z^*}.

Proof. For $n = 0$, the equivalence is immediate.

Assume that $\forall x(\tau_{\bar{k}}(x) \longleftrightarrow x = a_k)$ is true in L_{Z^*} for given k. Then $\tau_{Sc(\bar{n})}$ is the formula $\forall x(\tau_{\bar{n}}(y) \to \Sigma(x, y, a_1))$, which by the induction hypothesis applied to $\tau_{\bar{n}}$ is equivalent to $\forall y(y = a_n \to \Sigma(x, y, a_1))$, i.e., to $\Sigma(x, a_n, a_1)$. (B) now follows by (A).

C. If Φ is a primitive predicate of L_{Z^*}, then $[\Phi(\bar{n}_1, \ldots, \bar{n}_k)]^\tau$ is true in L_{Z^*} if and only if $\Phi^\tau(a_{n_1}, \ldots, a_{n_k})$ is true in L_{Z^*}.

Proof. Since Φ is primitive, $\Phi^\tau = \Phi$. $\Phi(\bar{n}_1, \ldots, \bar{n}_k)^\tau$ is then

$$\forall x_1 \ldots \forall x_k\left(\bigwedge_{j<k} \tau_{n_j}(x_j) \to \Phi(x_1, \ldots, x_k)\right),$$

which by (B) applied to $\tau_{n_1}, \ldots, \tau_{n_k}$ holds in L_{Z^*} iff $\Phi(a_{n_1}, \ldots, a_{n_k}) = \Phi^\tau(a_{n_1}, \ldots, a_{n_k})$ holds in L_{Z^*}.

(11) now follows from (C) by induction on the complexity of $\Phi(x_1, \ldots, x_k)$. Thus, we see that although the equivalences of form (11)

typically possess no first-order proof in L_{Z^*}, they may be shown to be true by reference to an independently given semantics for the atomic sentences of L_{Z^*}.

Having said all this, let us consider a substitutional construal of quantification in L_{ZF} which is adequate as regards satisfaction of the condition (D). We shall specify a class of individual constants denoting a certain countable model of ZF, the so-called 'minimal' model, which are susceptible of recursive translation into English. The minimal model results from restricting the field of the set-membership predicate to the class of those sets for which there is a predicate F of L_{ZF} such that

$$\mathrm{ZF} \vdash \exists y \forall x (x \in y \longleftrightarrow Fx).$$

This model is not elementarily equivalent to the usual one.[18] Thus, it should be clear at the outset what the interpretation to be suggested does *not* accomplish: it does not provide a substitutional construal of quantifiers that yields a distribution of truth-values to the sentences of L_{ZF} which is extensionally correct *vis-a-vis* the intended interpretation. What it does do is provide a substitutional interpretation of ZF which leads to an ω-consistent extension of ZF, one with respect to which ZF constitutes a correct description of a certain model, and so provides an explicit illustration of how Weston's problem may be avoided. For it is a familiar fact from model theory that if M is any model of a first-order theory T and T^* results from T by the addition of a set of individual constants denoting the elements of M, then the expansion of M to L_{T^*} associating each constant with its denotation in M is a model of T^*. Thus, let us fix an effective class $b_1, b_2, \ldots$ of individual constants, and let $F_1, F_2, \ldots$ be a recursive enumeration of those predicates of L_{ZF} for whose extension there is a proof of existence in ZF; each b_i will be construed as denoting the extension of F_i with respect to the normal interpretation. If ZF^* results from ZF by addition of the b_i, the atomic formulas of L_{ZF^*} are effectively interpretable by sentences of English as follows:

$$\ulcorner x \in b_i \urcorner \quad \text{by} \quad \ulcorner \exists y (\forall z (z \in y \longleftrightarrow F_i z) \;\&\; x \in y) \urcorner$$

$$\ulcorner b_i \in x \urcorner \quad \text{by} \quad \ulcorner \exists y (\forall z (z \in y \longleftrightarrow F_i z) \;\&\; y \in x) \urcorner$$

$$\ulcorner b_i \in b_j \urcorner \quad \text{by} \quad \ulcorner \exists y_1 \exists y_2 (\forall z_1 (z_1 \in y_1 \longleftrightarrow F_i z_1) \;\&\; \forall z_2 (z_2 \leftarrow y_2 \longleftrightarrow F_j z_2) \;\&\; y_1 \in y_2) \urcorner.$$

Thus, it seems we have a substitutional construal in quantifiers in ZF that satisfies Davidson's learnability stricture, and yet is based upon a substitution class harboring infinitely many individual constants.

III

The foregoing discussion attempted only to sustain the possibility of a substitutional interpretation of quantifiers in languages with infinitely many individual constants consistent with Davidson's constraint (D). Such an interpretation would allow us to provide an assignment of truth-conditions to all sentences of, for example, an extension of L of L_{ZF} given an assignment of truth-conditions to the atomic sentences of L. How is the latter to be specified? If we follow the previous example we would provide, in effect, that each atomic sentence of the form $\ulcorner b_i \in b_j \urcorner$ is true iff the extension of b_i is an element of the extension of b_j, where the extension of b_n is the class of all sets satisfying F_n. However, relative to the ontological purposes which usually motivate one to entertain a substitutional construal of quantifiers in this context, this interpretation is, of course, entirely self-defeating. For, the assignment of a referential interpretation to the atomic sentences would induce a referential interpretation of its quantifiers; the distinction of referential and substitutional quantification then loses its ontological significance. Accordingly, to achieve an ontologically non-trivial result, we must formulate a *non* referential analysis of the atomic sentences, i.e., we must specify an assignment of truth-conditions to the atomic sentences which does not accord a referential interpretation to the primitive predicates and singular terms of the language in question.

At first sight, this description of the enterprise may make it seem impossible; there simply seems to be no available form of semantic analysis for the primitive predicates of a first-order language other than the referential. And in one sense, this is quite true. However, one must distinguish two ways in which reference may play a role in providing truth-conditions for the atomic sentences: (a) we may use reference directly in interpreting the predicates and singular terms of the atomic sentences, according the former satisfaction conditions and the latter denotation conditions, or (b) we may appeal to reference in providing a finitely based assignment of truth-conditions for the class of atomic sentences *as a whole*, without attributing a referential function to their constituent predicates and singular terms. That (b) is strictly weaker than (a) may be seen by again considering our example of the language of virtual class theory as supplemented with substitutional quantification. Atomic sentences of the form $\ulcorner t \in \{x: Fx\}\urcorner$ may be assigned the truth-condition of the corresponding sentence $\ulcorner Ft\urcorner$. Each atomic sentence is then interpreted by another sentence, free of set-theoretic expressions. The atomic sentences are thus construed without ascribing a referential interpretation to the set-membership predicate; it functions merely as a place-holder, comparable to punctuation.

Let us now tie these observations together into a general picture. Suppose that **L** is a substitutional first-order language with substitution class C_L, and that $\mathbf{L}_0$ is a language for which a semantic description is given. A semantic description of **L** is now given in stages. First, the substitutional interpretation of quantifiers in **L** allows us to represent the truth-value of any sentence of **L** as a function of the truth-values of the atomic sentences of **L**. An effectively computable function ψ is then specified which maps the atomic sentences of **L** onto sentences of $\mathbf{L}_0$, and the truth-condition of any atomic sentence of **L** is identified with that of its counterpart in L_0.[19]

Now suppose that ξ is a model-theoretic interpretation of $\mathbf{L}_0$. We now extend ξ to an interpretation ξ^* of **L** as follows. Since ξ assigns a truth-value to every (closed) sentence of $\mathbf{L}_0$, we may use ψ to extend ξ to the atomic sentences A_L of **L**, thus:

$$(x)(x \in A_L \rightarrow (x \text{ is true} \longleftrightarrow \exists y(y = \psi(x) \text{ and } y \text{ is true})).$$

Suppose, then, that ξ^* assigns a (unique) truth-value to all sentences ϕ of **L** of complexity k or less. If ϕ is $\ulcorner A \vee B\urcorner$, where A and B are of complexity k or less, then $\xi^*(\ulcorner A \vee B\urcorner) = T$ if and only if $\xi^*(A) = T$ or $\xi^*(B) = T$, and similarly for the other truth-functional connectives. Finally, if $\phi = \ulcorner \forall \alpha A\alpha\urcorner$, where the substitution instances of A in C_L are of complexity k or less, then $\xi^*(\ulcorner \forall \alpha A\alpha\urcorner) = T$ if and only if $\forall \beta(\beta \in C_L \rightarrow \xi^*(\ulcorner A\alpha/\beta\urcorner) = T)$.

Any interpretation of ξ of $\mathbf{L}_0$, then, uniquely determines an interpretation ξ^* of **L**. (ξ^* is also defined on $\mathbf{L}_0$.) If S is a set of sentences of $\mathbf{L}_0$ or of **L**, we now provide that S is model-theoretically *consistent* if and only if some interpretation ξ of $\mathbf{L}_0$ is such that $\xi^*(\phi) = T$ for each $\phi \in S$. The definitions of validity and entailment are now derived from that of consistency in the usual manner.

Regarding the model theory of **L** as thus supervenient on that of $\mathbf{L}_0$ allows us to make precise our response to Weston's dilemma in Section II.

There will, indeed, typically be logical relations among the sentences of **L** which are not representable by means of first-order logic in **L**. However, this of itself is not a serious matter. For, the present approach to the model theory of **L** explains how the entailment relation in **L** may outrun those pairs $\langle \alpha, \beta \rangle$ of sentences that may be certified to be in that relation by virtue of a derivation of β's first-order schema from α's in quantification theory. It does this by ascribing to syntactically simple sentences of **L** the model-theoretic structure of sentences of $\mathbf{L}_0$, which need not be syntactically or semantically simple.[20] As regards ZF, however, all of this merely *sets* the problem, which is to specify an assignment of truth-conditions to the atomic sentences of (a substitutional extension of) $\mathbf{L}_{ZF}$ in a language which does not itself involve the very ontological commitments that the appeal to the substitutional construction is designed to avoid.

The Johns Hopkins University
The University of Michigan

NOTES

[1] 'A Plea for Substitutional Quantification', *The Journal of Philosophy* **LXVIII**, 8 (April 22, 1971), 231–237.
[2] 'Theories Whose Quantification Cannot Be Substitutional', *Nous* 8 (1974), 361–369.
[3] *Ontological Relativity and Other Essays*, Columbia University Press, New York, p. 106, 1969.
[4] 'A Plea for Substitutional Quantification', *op. cit.*, p. 234. Parsons is concerned with at most predicative set theory.
[5] *Op. cit.*
[6] 'A Plea for Substitutional Quantification', *op. cit.*, p. 233.
[7] 'Modalities and Intensional Languages', in Marx Wartofsky (ed.), *Studies in the Philosophy of Science*, Reidel, Dordrecht, pp. 77–96, 1963.
[8] 'A Plea for Substitutional Quantification', *op. cit.*, p. 233.
[9] Dunn and Belnap in 'The Substitutional Interpretation of the Quantifiers', *Nous* ii (1968), 184, defend the possibility of using substitutional quantification in the metalanguage. For a criticism of their view, see S. A. Kripke, 'Is There a Problem About Substitutional Quantification?' in Evans and McDowell (eds.), *Truth and Meaning*, Oxford, pp. 341–342, 1976.
[10] Specifically, we may define via well-known methods predicates 'Cx' and 'Dxy' expressing 'x is (the Gödel number of) a closed singular term' and 'y is the denotation of the term (with Gödel number) x', respectively, and a predicate 'sub(x, y)' expressing 'x is (the Gödel number of) a predicate and y is (the Gödel number of) a substitution instance of x' for each extension E of ZF meeting Weston's conditions. We now

add to L_E the predicate 'True', construed as applying to Gödel numbers of formulae of L_{ZF}, and add to E (i) each instance of the schema '$\phi \rightarrow \text{True}(\ulcorner \phi \urcorner)$', wherein $\ulcorner \phi \urcorner$ represents (the Gödel number of) a sentence of L_{ZF}, and (ii) each instance of schema '$(x)(y)(Cx \,\&\, Dxy \rightarrow Fy) \rightarrow (z)(\text{sub}(\ulcorner F \urcorner, z) \rightarrow \text{True}(z))$' for each predicate '$Fx$' of L_{ZF}. It is clear that the axioms obtained through (i) and (ii) are correct with respect to the substitutional as well as the referential interpretation of the quantifier.

[11] It may be shown that the axioms of pairing and finite unions suffice, given our other assumptions, to insure this sort of representability.

[12] Weston, *op. cit.*, p. 362.

[13] More precisely, if S is a set of semantical rules generating truth-conditions for the sentences of L, and e is an expression of L, we say that e is *semantically composite* relative to S iff any subset of S providing truth-conditions for all sentences of L not containing e suffices to provide truth-conditions for all sentences which do contain e (i.e., for all sentences of L). We now characterize as semantically primitive with respect to S any expression of L which is not semantically composite with respect to S.

[14] This is to say, some recursive function acting on the sentences of L yields a truth-condition for each sentence of L in English. Weston's description of Davidson's requirement, reformulated slightly (his condition [5]), is as follows:

> [5] A translation F defined on L is satisfactory only if there is a finite subset S of singular terms and function symbols of L and a recursive function G acting on finite sequences from $F(S)$ such that for each singular term t one can effectively find a finite sequence s of members of $F(S)$ such that $F(t) = G(s)$.

It is obvious that any translation F satisfying [5] is effective. It may be less obvious that any effective F satisfies [5] as well. For, let F be effective, acting on expressions of L, and let t be a singular term of L. Suppose that f recursively enumerates the range of F. Letting $S = \{F(t)\}$ in [5], for any singular term r of L, let $f(n) = F(r)$ and s_n be the sequence of elements of S of length n. Then $F(r) = f(\text{length}(s_n))$, and F satisfies [5]. Thus [5] adds nothing to the constraint that a satisfactory translation must be effective, and so we shall speak simply of effectiveness, and ignore the more cumbersome formulation of [5].

[15] p. 364.

[16] That is, for any formula ϕ of L_Z, $Z \vdash \phi$ only if $Z^* \vdash \phi^T$.

[17] Say, $\forall x[\psi_1(x) \longleftrightarrow \forall y(y \in x \longleftrightarrow \forall z(z \in y \longleftrightarrow z \neq z))]$.

[18] This is most easily seen by considering the ZF-translation of the statement '$(Ex)(x$ is a model of ZF$)$'. This sentence is true with respect to the standard model, but false with respect to the minimal model. We are indebted to an anonymous referee for this example.

[19] The latter need not, of course, be atomic.

[20] In this we differ from S. A. Kripke, 'Is There a Problem About Substitutional Quantification?' in Evans and McDowell (eds.), *Truth and Meaning: Essays in Semantics*, Oxford, § 1(a), 1976. Kripke offers a construction that is similar to ours in providing for a hierarchy of L over L_0 for the purposes of semantics, but then ignores the L_0 structure imposed on the atomic sentences in characterizing validity. This seems to us unmotivated.