THE JOURNAL OF SYMBOLIC LOGIC
Volume 61, Number 2, June 1996

DID TARSKI COMMIT "TARSKI'S FALLACY"?

G. Y. SHER

§1. In his 1936 paper, *On the Concept of Logical Consequence,* Tarski introduced the celebrated definition of **logical consequence**: "*The sentence* σ follows logically *from the sentences of the class* Γ *if and only if every model of the class* Γ *is also a model of the sentence* σ." [55, p. 417][1] This definition, Tarski said, is based on two very basic intuitions, "essential for the proper concept of consequence" [55, p. 415] and reflecting common linguistic usage: "Consider any class Γ of sentences and a sentence which follows from the sentences of this class. From an intuitive standpoint it can never happen that both the class Γ consists only of true sentences and the sentence σ is false. Moreover, ... we are concerned here with the concept of logical, i.e., *formal,* consequence." [55, p. 414] Tarski believed his definition of logical consequence captured the intuitive notion: "It seems to me that everyone who understands the content of the above definition must admit that it agrees quite well with common usage. ... In particular, it can be proved, on the basis of this definition, that every consequence of true sentences must be true." [55, p. 417] The formality of Tarskian consequences can also be proven. Tarski's definition of logical consequence had a key role in the development of the model-theoretic semantics of modern logic and has stayed at its center ever since. [2]

In a recent book, **The concept of logical consequence** [9], J. Etchemendy has launched an all out attack on Tarski's definition: "[M]y claim is that Tarski's analysis is wrong, that his account of logical truth and logical consequence does not capture, or even come close to capturing, any pretheoretic conception of the logical properties." [9, p. 6] "This book consists of a single, extended argument. The conclusion of the argument is that the standard, semantic account of logical consequence is mistaken." [9, p. 8] "Various characteristics distinguish logical truths

Received January 30, 1995; revised June 7, 1995.

I have benefited from many discussions on issues related to this paper. I would like to thank, in particular, Jon Barwise, Patricia Blanchette, George Boolos, Richard Cartwright, John Etchemendy, Allen Hazen, Philip Kitcher, Charles Parsons, Joe Ramsey, Stephen Read, Oron Shagrir, Stewart Shapiro, Peter Sher, Wilfried Sieg and the students in my class "What is Logic?". Special thanks go to Wilfrid Hodges for his comments and correspondence. I would also like to thank an anonymous referee of the JSL. Early notes leading to this paper were presented to the Logic Colloquium, UC Berkeley and to the Eastern APA meeting, 1991. Two lectures based on the penultimate version were delivered to the Philosophy of Mathematics Workshop, UCLA. I wish to thank the participants for stimulating comments.

[1] To unify the symbolization throughout the paper, I replace "X" and "K" in the citations from Tarski by "σ" and "Γ", respectively. (Henceforth I will abbreviate "if and only if" as "iff".)

[2] (a) In this paper I disregard the differences (if any) between Tarski's 1936 definition and the ones commonly used today, including any differences between Tarski's 1936 notion of model and the modern notion. (b) My use of syntactic devices to distinguish use and mention is casual.

0022-4812/96/6102-0016/$4.40

from common, run-of-the mill truths, and logically valid arguments from those that happen to have a false premise or a true conclusion. But Tarski's analysis does not capture any of these characteristics ... Furthermore, we are not even guaranteed that [Tarski's] definition will be extensionally correct when applied to a given language, not even in the paradigmatic, first-order case." [9, p. 136] Tarski, according to Etchemendy, fell prey to a series of unfortunate confusions, one of which is (if Etchemendy is right) "Tarski's fallacy": Tarski made a simple, elementary mistake in calculating modalities, and this led him to assert that his definition captured the intuitive notion. In fact, Tarski's definition is a definition of material, not logical, consequence, though this may have eluded the logical and philosophical communities.

Etchemendy's verdict is rather extreme, and if founded, its impact on our assessment of contemporary logic, not to speak of Tarski himself, will be profound. Is his verdict justified? I will begin with Tarski's pretheoretic account. One way to interpret Tarski's first intuitive consideration on logical consequence is as follows: Assume σ is a logical consequence of Γ. Then *it is impossible* that all the sentences of Γ are true and σ is false (where "it is impossible" is an intuitive modal operator, equivalent to "necessarily, it is not that case that"). Thus interpreted, Tarski's pretheoretic notion of logical consequence involves two intuitive ideas: the idea that *logical consequence is necessary* and the idea that *logical consequence is formal.* These ideas play the role of adequacy conditions: an adequate definition of logical consequence yields only consequences that are necessary and formal. Leaving formality aside, what Tarski had to do to justify his definition is, then, prove

(1) $$\Gamma \models \sigma \rightarrow \Box(\text{All the sentences of } \Gamma \text{ are true} \rightarrow \sigma \text{ is true})$$

where $\models$ is the relation of Tarskian consequence and $\Box$ is an intuitive necessity operator. Tarski, as we have seen, claimed the first adequacy condition was provable from his definition, but he did not indicate what the proof was. Etchemendy [9, p. 86] takes Tarski's proof to have the following structure: Assume

(i) $\Gamma \models \sigma$,

(ii) $\neg$(All the sentences of Γ are true $\rightarrow$ σ is true), or equivalently: All the sentences of Γ are (actually) true & σ is (actually) false.

These assumptions are contradictory: (i) says that σ is true in every model in which all the sentences of Γ are true, but (ii) says that there is a model, namely, one representing the "actual world", in which all the sentences of Γ are true and σ is false. However, Etchemendy points out, what a proof with this structure proves is

(2) $$\Box[\Gamma \models \sigma \rightarrow (\text{All the sentences of } \Gamma \text{ are true} \rightarrow \sigma \text{ is true})].$$

And (2) does not imply (1). In Etchemendy's words: "To show that all Tarskian consequences are consequences in the ordinary sense, we would need to prove a theorem with embedded modality. ... Obviously, the proof in question does not show that every Tarskian consequence is a consequence 'in the ordinary sense.' It is only through an illicit shift in the position of the modality that we can imagine ourselves demonstrating *of* any Tarskian consequence *that* it is entailed by [i.e., follows with necessity from] the corresponding set of sentences." [9, p. 87–8] To take Tarski's proof as a proof of (1) is, then, to make the fallacious inference:

$\Box(\mathscr{P} \rightarrow \mathscr{Q})$; therefore, $\mathscr{P} \rightarrow \Box\mathscr{Q}$. This is *Tarski's fallacy.* Tarski made a basic mistake in working out his modalities, and coming generations of logicians could have been (have been?) misled into believing that model-theoretic semantics is sound.

Did Tarski commit "Tarski's fallacy"? If we take this question as a historical question, the answer is quite simple: Tarski declared that the first intuitive condition was provable from his definition, but he did not specify (or indicate in any way) what the proof was. What Etchemendy takes Tarski's proof to be is, therefore, based on speculation. It is consistent with a certain conception of Tarskian semantics (the "interpretational" conception, which I will discuss below), but aside from that, Etchemendy does not present any piece of evidence that would connect this proof with Tarski. Furthermore, the alleged fallacy is not so much a modal fallacy, as a fallacy in handling a narrow scope operator. To prove the statement: $\mathscr{P} \rightarrow \mathscr{F}(\mathscr{Q})$ – where $\mathscr{F}$ is a sentential operator – Tarski, according to Etchemendy, assumed $\mathscr{P}$ and derived a contradiction from $\neg\mathscr{Q}$. It is hard to believe that any competent logician would give this kind of proof (unless, of course, in this particular case, $\mathscr{Q}$ implied $\mathscr{F}(\mathscr{Q})$).

Etchemendy's criticism, however, directs our attention to the fact that Tarski never (publicly) proved the adequacy of his definition. In this paper I would like to examine the adequacy of Tarski's definition with respect to modern logic. More precisely, I am interested in whether modern logic has the resources for developing an adequate account of logical consequence in the spirit of Tarski. Although my interest is not historical, I find Tarski's paper on logical consequence useful in clarifying the motivation and key concepts of the semantic project and I will keep turning to it throughout the investigation. To fix Tarski's 1936 [55] paper as a point of reference, I will give a short synopsis of this paper.

Tarski opened his paper with a statement regarding the use of "logical consequence" in mathematics and logic. The meta-mathematical use of "consequence" is not arcane, according to Tarski, but aims to simulate the *ordinary* use of the term. Tarski pointed out the insufficiency of the proof-theoretic definition, and proceeded to lay down the two intuitive conditions on an adequate definition of logical consequence. Tarski then considered a substitutional definition of logical consequence, but found the substitutional definition inadequate. At this point Tarski declared that we must look for a new (non-syntactic) means for expressing the intuitive conditions, suggesting that "[s]uch a means is provided by semantics." [55, p. 416] In preparation for his semantic definition Tarski introduced two concepts: the concept of *satisfaction,* defined in *The Concept of Truth in Formalized Languages,* and the concept of *model,* which he defined in terms of satisfaction. The definition of logical consequence in terms of model followed, accompanied by the claim that the two pretheoretic conditions were satisfied. Up to this point the context of Tarski's investigations appears to be that of Russellian type-theoretic logic (with simple types).[3] But in the conclusion to his paper Tarski expanded the context of his inquiry, observing that underlying his "whole [semantic] construction is the division

[3]Tarski did not explicitly indicate what logic he had in mind, but a look at his cross references (e.g., the reference to Tarski [53] in fn. 1, p. 410) as well as a comparison with other articles from the same period indicate that this was the notion of logic he assumed in the 1936 paper.

of all terms of the language discussed into logical and extra-logical." [55, p. 418] This observation put the definition of logical consequence in a new light: It is possible that for some selections of logical terms the definition yields the right results whereas for others it does not. (Indeed, it is easy to see that for some selections of logical terms Tarski's definition fails.) Tarski said he did not know whether a systematic distinction between logical and non-logical terms was forthcoming, and with this uncertain note he ended his paper.

Does the failure of Tarski's definition for some selections of logical terms undermine his claim of having proved its adequacy? I think it does not. Tarski formulated his definition for the standard logic (or one of the standard logics) of his day, and this logic admits a specific selection of logical terms. Tarski's justification was intended therefore for this specific selection. What the lesson at the end of his paper indicates is that the standard selection, or certain features of the standard selection, plays an essential role in an adequate justification.

Back to modern logic: One natural way of going about the proof of (1), where $\models$ is relativized to a (standard) language $\mathfrak{L}$,[4] is as follows: assume

(3) $$\Gamma \models \sigma,$$

assume

(4) $$\diamond(\text{All the sentences of } \Gamma \text{ are true } \& \ \sigma \text{ is false}),$$

where $\diamond$ is an intuitive possibility operator, and derive a contradiction. Unpacking assumption (3), we get: *There is no* model for $\mathfrak{L}$ in which all the sentences of Γ are true and σ is false. If we can show that (4) implies that *there* is a model for $\mathfrak{L}$ in which all the sentences of Γ are true and σ is false, we will have derived the desired contradiction. To prove (1) it is thus sufficient to show that given a set Σ of sentences of $\mathfrak{L}$ (in our case, $\Sigma = \Gamma \cup \{\neg\sigma\}$), the following holds:

(5) If it is intuitively possible that all the sentences of Σ are true, then there is a Tarskian model (a model for $\mathfrak{L}$) that realizes this possibility.

Does (5) hold in standard semantics?

The answer to this question depends, in part, on our understanding of the intuitive concepts involved. But at least in one case, that of standard first-order logic, the vindication of (1) appears to be straightforward: Consider the logical axioms and rules of proof of a standard first-order system, $\mathscr{L}$, with a language $\mathfrak{L}$. Examining the axioms and rules of proof of $\mathscr{L}$ one by one, we intuitively verify that all the axioms are necessarily true and all the rules of proof are necessarily truth preserving. We conclude that the axioms and rules of proof of $\mathscr{L}$ satisfy the necessity condition. Now, take any set Σ of sentences of $\mathfrak{L}$. Assume it is intuitively possible that all the sentences of Σ are true. Then, because the axioms and rules of proof in $\mathscr{L}$ satisfy necessity, no contradiction can be derived from Σ in $\mathscr{L}$, i.e., Σ is proof-theoretically

[4] In this paper, "standard language" and "standard logic" refer to a language and a system of logic as in standard textbooks of mathematical logic, e.g., Enderton [7]. Often, but not always, the point I wish to emphasize is the restriction of logical terms to the "standard" ones: truth functional connectives, the universal and/or existential quantifier and identity. "Standard logic" also includes "standard" higher-order logic (with "standard" semantics). See, e.g., Shapiro [45].

consistent. It follows from the completeness theorem (Gödel-Henkin-Maltsev) that Σ has a (Tarskian) model. Q.E.D.[5]

This proof is short and elegant, yet it does not go far enough. First, Tarski's definition of logical consequence applies to standard higher-order logic as much as to standard first-order logic, but the argument from completeness does not apply to standard higher-order logic. Second, even with respect to standard first-order logic we may seek a level of explanation that goes beyond this proof. The situation is somewhat similar to that of using completeness (Henkin's proof) to establish the Löwenheim-Skolem theorem. We would like to understand how uncountable models of consistent theories are contracted to countable models (how, for example, the uncountable membership relation of a "large" model of ZF is reduced to a countable relation), but the proof from completeness does not help us to obtain this understanding. In the present case, we would like to know how models represent all intuitive possibilities with respect to a given language: what features of possible states of affairs correlate with what features of models, how differences between models suffice to represent *all* relevant differences between possible states of affairs, etc. But the proof from completeness does not illuminate these issues. Furthermore, in view of Tarski's qualifying note, we would like to understand what role logical terms play in the correspondence between models and possibilities. The proof from completeness relies on the connection between the standard selection of logical terms and the standard proof method, but it does not explain what role this selection plays in shaping the model-theoretic apparatus.

§2. Let us call the claim that Tarski's definition satisfies the intuitive conditions of formality and necessity "Tarski's adequacy conjecture." Tarski's adequacy conjecture is an informal conjecture, and our verdict on it may vary according to our understanding of the intuitive and technical notions involved. The extension of "logical term" is an important factor but not the only factor. The technical notion of model, the intuitive or philosophical notions of necessity (possibility) and formality, and perhaps other notions as well, play a crucial role. We can regard Tarski's definition as a schema that for different notions of logical system (logical term, model, etc.) yields different definitions and ask: For what kind of logical system and what notions of necessity and formality does Tarski's adequacy conjecture hold? Our question, then, concerns conceptions of logical semantics, where by "conception of logical semantics" I mean a view, or a theory, of

(i) the technical apparatus of logical semantics,
(ii) the intuitive (or philosophical) notion of logical consequence, and
(iii) the relation between the two (including the philosophical underpinning of (i)).

[5]A similar proof is given in Kreisel [20], though in a somewhat different context. Assuming the intuitive adequacy of the standard definition when "every model" ranges over structures involving classes in general, Kreisel asks whether the definition preserves its adequacy when models are restricted to proper sets. Using completeness, Kreisel shows that at least in the case of standard first-order logic the answer is positive. Etchemendy [9, Ch. 11]) claims Kreisel's proof "does not count" because of the assumption mentioned above. However, the proof formulated in the text above does not involve this assumption.

In positive terms our question is: Is there a conception of logic, or a strategy for constructing logic, which vindicates Tarski's conjecture in a non-trivial manner?

Etchemendy's claim that Tarski's definition fails to capture the intended notion of logical consequence is relative to a specific conception of logical semantics, the so-called *interpretational* conception. And relative to this conception, his claim is correct: If modern semantics is an interpretational system, then indeed Tarski's conjecture fails for this semantics. I will show, however, that modern semantics is not an interpretational system, that, furthermore, the interpretational strategy is an unreasonable strategy for constructing a logical semantics. In the present section I will analyze the two conceptions of logic entertained by Etchemendy as strategies for constructing a logical semantics. In Section 3, I will propose a third conception which, I will claim, vindicates Tarski's conjecture.

In presenting the two conceptions of logic below I will not follow Etchemendy's accounts to the letter. Instead, I will describe two broad approaches to the problem of formulating an adequate account of logical consequence, suggesting a motivation and analyzing the central difficulties involved.[6]

The metaphysical (representational, inflationary) conception[7]**.** The metaphysical conception of logical semantics conflates the notion of logical consequence with that of necessary consequence in general. One motivation for constructing such a semantics is based on (i) the central place of necessity in the intuitive characterization of logical consequence, and (ii) the difficulty in identifying the exact difference between logical and necessary consequence, i.e., the precise content of formality. Etchemendy expresses a common view when he says: "The most important feature of logical consequence, as we ordinarily understand it, is a modal relation that holds between implying sentences and sentence implied. The premises of a logically valid argument cannot be true if the conclusion is false; such conclusions are said to 'follow necessarily' from their premises. That this is the single most important feature of the consequence relation, or at any rate of our ordinary understanding of that relation, is clear from even the most cursory survey of texts on the subject." [9, p. 81] [8] But while most people would not identify the notions of logical and necessary consequence, drawing a sharp distinction between the two is not simple, and the attempts to account for the distinctive features of logical consequence have often taken too much for granted. The most influential approach, due to Quine, maintains that genuinely logical consequences are due to the *logical structure* of sentences, where logical structure is a function of the specific logical terms and their arrangement in a given sentence: "Logical implication rests wholly on how the truth functions, quantifiers, and variables stack up. It rests wholly on what we may call, in a word, the logical structure of the ... sentences [involved]." [39, p. 48]

[6]My interest in these conceptions is purely theoretical. With one exception, I am not claiming that anyone actually accepts either of these conceptions. That exception is Etchemendy who, in spite of certain tensions he is aware of (see Ch. 5 and Ch. 8, pp. 112–24), views modern logical semantics as an interpretational theory. (See [9, p. 51] and note the references to the standard semantics of mathematical logic as "the standard interpretational semantics" on pp. 112, 122, 126, etc.) For Etchemendy's motivation see Section 4 below.

[7]I prefer "metaphysical" to "representational" since in each of the three conceptions I will discuss models represent something: it is what they represent, rather than whether they represent, that (among other things) distinguishes between them.

[8]See references to other writers who share this view on pp. 81–2.

But underlying this account is a certain view of *logical structure* and its carriers, and Quine, as well as others who espoused this view, did not give an adequate explanation of why the *logical* structure of language rests on certain terms rather than on others.[9] (Indeed, few philosophers have offered systematic grounds for the distinction between logical and non-logical terms, and even fewer have done so from a semantic perspective.[10])

One way to avoid this and other related problems is to take the broader category of necessary consequence as basic and the category of logical consequence as derivative upon it: say, logical consequences are the most robust necessary consequences, where robustness is measured by pragmatic, behavioristic, or other criteria. How do we construct a system for identifying necessary consequences? The idea of *metaphysical* or *representational* semantics is that of a dual system whose constituents are

(i) a fully interpreted language and
(ii) an apparatus of models representing possible worlds, i.e., possible variations in the actual world that would affect the truth value of sentences of the language. The two are connected by
(iii) a definition of truth (satisfaction) in a model.

Etchemendy describes the guidelines for models and truth in a model as follows: "the class of models should contain representatives of *all* and *only* intuitively possible configurations of the world" [9, p. 23]; "a sentence is to be *true in* a model iff it *would have been true* had ... the world actually been as depicted by the model." [9, p. 24] The definition of necessary consequence is Tarski's: given a language $\mathfrak{L}$, σ is a necessary consequence of Γ iff there is no model for $\mathfrak{L}$ (possible world with respect to $\mathfrak{L}$) in which all the sentences of Γ are true and σ is false. Clearly, consequences falling under this definition (i.e., consequences preserving truth in all possible worlds) satisfy the intuitive condition of necessity.

Is the standard semantics of mathematical logic a metaphysical semantics? Is the metaphysical strategy for constructing a logical semantics worth pursuing? In spite of some superficial similarities it is unreasonable to view the standard semantics of mathematical logic as a metaphysical (representational) semantics. First, the language of mathematical logic is only partly interpreted (only the logical terms are fully interpreted), but the language of metaphysical semantics is fully interpreted (the notion of necessary consequence in general requires a fully interpreted language). Second, while the model theory of mathematical logic is couched in a background theory based on set theory, the model theory of metaphysical semantics is couched in a background theory based on general metaphysics. In particular, the constraints on metaphysical models are essentially more involved than those on logical models, and they rule many models of mathematical logic illegitimate.

---

[9]In **Philosophy of logic**, Quine justifies his preference of standard first-order logic with its specific selection of logical terms on the ground that this system is complete. But note that some systems with non-standard logical terms, e.g., first-order logic with the quantifier "there exist uncountably many," are also complete. (See Keisler [19].) In **Word and object** [37, pp. 57–61], Quine proposes a *partial* behavioristic justification for the standard logical terms.

[10]See Peacocke [34] and McCarthy [28]. There are certain similarities between Peacocke's and McCarthy's approaches and mine, but by and large my account will be different from theirs.

Thus, take a simple interpreted language, $\mathfrak{L}$, and consider the following three inference schemata of $\mathfrak{L}$:

(6) $\mathfrak{a}$ is yellow (all yellow); therefore, $\mathfrak{a}$ is not red,

(7) $\mathfrak{a}$ is yellow; therefore, $\mathfrak{a}$ is not round,

and

(8) $\mathfrak{a}$ is round; therefore, $\mathfrak{a}$ is not square,

where $\mathfrak{a}$ is to be replaced by an individual term of $\mathfrak{L}$. To distinguish necessary from non-necessary inferences of $\mathfrak{L}$, the apparatus of models for $\mathfrak{L}$ has to be constructed in such a way that all instances of (6) and (8) preserve truth in all models while some instances of (7) do not. I.e., the meta-theory has to include instructions that will produce the same results as:

(a) the extension of "x is yellow" shall form an empty intersection with the extension of "x is red" in all models;
(b) the extension of "x is round" shall form an empty intersection with the extension of "x is square" in all models;
(c) the extension of "x is yellow" shall form a nonempty intersection with the extension of "x is round" in some models.

More generally, the meta-theory has to include information about all the possible and necessary interrelations between the individuals, properties and relations denoted by terms of the given language. These instructions will rule out many models that standard mathematical logic assigns to (its codification of) $\mathfrak{L}$: models in which the intersection of the extensions of "x is yellow" and "x is red" is not empty, models in which the intersection of the extensions of "x is round" and "x is square" is not empty, etc. While mathematical logic imposes primarily set-theoretic constraints on its apparatus of models (e.g., if in a given model the intersection of two extensions is not empty, then in that model each of the two extensions is not empty), metaphysical logic imposes (in addition to these) a great many metaphysical constraints.

This difference between logical and metaphysical semantics points to deep problems that make the construction of metaphysical (representational) semantics unfeasible. First, the amount of information that has to be taken into account in constructing an apparatus of models for a reasonably rich language is enormous. It is doubtful that this information can be organized into anything like axiomatic set theory or any other manageable theory. Second, metaphysical semantics requires solutions to the most obscure and thorny questions of general metaphysics. Consider a discourse about a certain object, say my dog Cerberus. What are all the possible variations in the world vis-a-vis Cerberus? Clearly, there are models in which Cerberus is your dog rather than mine, but is there a model in which Cerberus is human? a model in which Cerberus is another dog? a model in which Cerberus evolved from an egg fertilized by a spermatozoon with an X chromosome? a model in which Cerberus was not begotten by his sire (Parker Jameson of Sunny-Duke)? a model in which Cerberus (this Cerberus) is a mythological beast? a model in which Cerberus is New York City? a model in which Cerberus is the empty set? Where does Cerberus stand with respect to "x is moral" and "x is rational"? (Is there a model in which Cerberus is a moral dog? a rational dog?) Here we come upon recalcitrant questions of identity, essential properties, moral and rational agency, meaning, etc. that have baffled philosophers for years.[11] (Axiomatic set theory also leaves many questions unanswered, and we know that in principle set theory is incomplete. But the "incompleteness" of general metaphysics is of a different order.) Furthermore, it is not even clear that we have a single, coherent notion of necessity (possibility), rather than a medley of vague and possibly incompatible notions.

In sum: (a) The semantics of modern mathematical logic is not a metaphysical (representational) semantics. (b) The strategy of vindicating Tarski's conjecture by developing a metaphysical semantics is, given the daunting problems facing general metaphysics, not feasible. We may try to salvage this strategy by embedding our account of logical consequence in a general theory of knowledge or meaning, but the obstacles facing us would be no less formidable. Describing logical semantics as a specific subsection of metaphysical semantics would not do either, since this is tantamount to giving an account of the distinctive nature of *logical* semantics.

The linguistic (substitutional/interpretational, deflationary) conception. How can we obtain a notion of consequence satisfying necessity and formality without resorting to an obscure possible world discourse? One strategy we may contemplate is this: instead of defining logical consequence in terms of variations in the world, let us *hold the world fixed* and define logical consequence in terms of *variations in language*. For example, to prove the logical *invalidity* of

(9) Proust is a novelist; therefore, Proust is not a dog

we do not have to produce a possible world in which some dogs are novelists; it suffices to produce a language in which "dog" plays the lexical role of "human," while "Proust," "novelist" and "not" preserve their usual lexical roles. Such a language will constitute a counterexample to (9), since the *actual* truth values of its premise and conclusion under this linguistic transformation are, respectively, T and F.[12] This linguistic method can also be used to establish the *validity* of inferences: for example, we can prove the validity of

(10) Some novelists are dogs; therefore, some dogs are novelists

by showing that no matter how we vary the lexical role of "dog" and "novelist" (subject to certain grammatical constraints), the resulting inference will not have a (materially) true premise and a (materially) false conclusion. I will call this strategy "the linguistic strategy" and the conception of logical semantics based on it "the

[11] A "classic" reference for some of these questions is Kripke [21]. See also Forbes [10]. Modal logic is not interested in the modal status of inferences like those in the examples above, but semantic considerations on modal logic show that allowing the language to include intensional operators will lead to additional deviations from standard logical semantics.

[12] In this paper I take "actual" and "material" to be synonymous. "Actual" or "material truth" is "truth in the actual world" and the notion of "actual world" is taken to be understood in advance. I do not need, nor do I want to commit myself to, a precise notion of "actual world." All I require is that we have a rough, intuitive notion of the actual world with some paradigm examples of what does and what does not fall under it (e.g., familiar, middle sized physical objects like bridges and persons do, while fictional characters and at least some "large" sets do not).

linguistic conception." The principle underlying the definition of logical consequence in *linguistic semantics* (semantic theories based on the linguistic conception) is the following: Given an interpreted language $\mathfrak{L}$, a sentence σ of $\mathfrak{L}$ and a set, Γ, of sentences of $\mathfrak{L}$, σ is a logical consequence of Γ (in $\mathfrak{L}$) iff under no permissible lexical variation in $\mathfrak{L}$ (using the lexical resources of any language $\mathfrak{L}' \supseteq \mathfrak{L}$) are all the sentences of Γ (materially) true and σ (materially) false. (The permissibility conditions have to do with uniformity and preservation of grammatical categories.)

Linguistic semantics reduces the notion of logical consequence to two tractable notions: (i) the notion of material truth, and (ii) the notion of lexical variations in language. The linguistic strategy, if successful, is highly attractive: avoiding metaphysical speculation altogether, it produces an ontologically cheap solution to a deep philosophical problem. However, the great simplicity of linguistic semantics calls for careful scrutiny: Is the reduction of the strong, modal notion of logical consequence to two relatively weak notions genuine? How does linguistic semantics prevent intuitively material consequences from satisfying the definition of logical consequence? I will examine two concrete examples of linguistic semantics: *substitutional semantics* and *interpretational semantics.*

Substitutional semantics. The substitutional method has often been conceived as a means for avoiding undesirable ontological commitments. For example, by defining the standard quantifiers substitutionally rather than objectually we can apply the rules of inference of first-order mathematical logic in fictional and mathematical discourse without committing ourselves to an ontology of nonexistent physical objects and abstract, mathematical objects.[13] To define logical consequence, we start with a single interpreted language, $\mathfrak{L}$, and we define: σ is a logical consequence of Γ iff there is no permissible (uniform) substitution of primitive terms of $\mathfrak{L}$ for primitive nonlogical terms of $\mathfrak{L}$ under which all the sentences of Γ are (materially) true and σ is (materially) false. Given a primitive term, t, of $\mathfrak{L}$, the terms substitutable for t in $\mathfrak{L}$ form a subclass of terms of $\mathfrak{L}$, the *substitution class* of t in $\mathfrak{L}$; membership in this class is based on *grammatical compatibility*: roughly, t is *grammatically compatible* with t′ iff replacement of t by t′ (and t′ by t) within any grammatical context results in a grammatical context; a substitution of t′ for t is *permissible* only if t′ is a member of the substitution class of t in $\mathfrak{L}$.[14]

Regrettably, however, the substitutional test of logical consequence is too weak. I will point out three ways in which it fails – three "fallacies of material consequence."

The first fallacy of material consequence. The first fallacy results from an unwarranted simplification of the linguistic conception. The test for logical consequence in the linguistic conception (i.e., preservation of truth under variations in language (lexicon)) involves a multiplicity of languages (lexicons). The substitutional test, however, is carried out within a single language. Given a language, $\mathfrak{L}$, substitutional semantics tests consequences for preservation of truth under replacements of

[13] See, for example, Marcus [25, 26] and Gottlieb [14]. But note fn. 14 below.

[14] My account of the substitutional test for logical consequence differs from Bolzano's [3] in that the latter involves variations in *ideas* rather than *words* and, related to this, a *nonmaterial* notion of truth. This difference means that Bolzano's substitutional semantics does not fall under the rubric of linguistic semantics and it cannot be viewed as an antidote to the metaphysical approach. In this paper I examine the substitutional approach only as such an antidote. Although Etchemendy talks about "Bolzano's definition," he has in mind a linguistic version of the substitutional definition. (See [9, fn. 2, pp. 162–3].)

terms of $\mathfrak{L}$ by terms of $\mathfrak{L}$, not under replacements of terms of $\mathfrak{L}$ by terms of other languages. In particular, if t is not a term of $\mathfrak{L}$, no substitution of t for terms of $\mathfrak{L}$ will ever be considered. The substitutional notion of logical consequence is, thus, relative to the lexical riches of a given language. It is easy to see that in some cases of lexically deficient languages the substitutional test fails. Thus, let the classes of primitive singular terms and 1-place predicates of $\mathfrak{L}$ be {"Tarski," "Łukasiewicz"} and {"x is Polish," "x is a logician"}. Then

(11) Tarski is Polish; therefore, Tarski is a logician

passes the substitutional test.

The second fallacy of material consequence. In discussing the substitutional theory of quantifiers Quine [38, p. 106] remarked that substitutional quantification makes good sense – the same good sense – no matter what terms the substitutional variables stand for. Even the left-hand parenthesis, Quine emphasized (using an example due to Lésniewski), can make up a substitution class for these variables. A similar principle applies to the substitutional definition of consequence. The substitutional method sets no constraints on what terms can be substituted for, and as long as the choice of substitutional classes obeys the grammatical requirement of compatibility anything is permitted. From the point of view of substitutional semantics, then, the standard distinction between logical and non-logical terms is arbitrary. The substitutional test is relative to arbitrary selections of *fixed* (*non-fixed*) terms, and it is easy to see that under some selections of fixed terms intuitively nonlogical inferences pass the test. Thus, let "Tarski" and "is a logician" serve as fixed terms: regardless of the lexical resources of the given language, (11) will come out logically valid. (The plight of fixed terms goes also in the other direction: if we take the material conditional as a non-fixed term, the intuitively logically valid rule of Modus Ponens will turn out invalid.)

The third fallacy of material consequence. Substitutional semantics investigates changes in the truth value of sentences under variations in language (variations in linguistic terms), not under variations in the world (variations in existent objects and their properties), and this enables it to embed its test of logical consequence in a material theory of truth. In particular, the notion of "truth under substitutions" is a material notion: a sentence is true under a given substitution iff another sentence (that obtained from it by the said substitution) is materially true.[15] Now, consider the inference

(12) Tarski is Polish; therefore, Kotarbiński is Polish

symbolized in a language, $\mathfrak{L}$, as: $P^1 t$; therefore $P^1 k$. Assume that terms in $\mathfrak{L}$ are divided into fixed and non-fixed in the standard way. (12) passes the substitutional test iff

(13) $$P^1 t \rightarrow P^1 k$$

[15]As I indicated above, I am only interested here in the substitutional method as used in conjunction with a material theory of truth. There may be, of course, other uses of the substitutional method. For example, we can understand Marcus' use of substitutional quantifiers in [25, 26] as meant to emphasize the ontological *neutrality*, rather than economy, of logic. (On the uses and limits of substitutional quantification see Parsons [32] and Kripke [22].)

is true under all permissible substitutions for P^1, t and k. I.e., (12) is substitutionally valid iff the meta-theoretical substitutional quantification

$$(14) \qquad (\forall P^1)(\forall x)(\forall y)(P^1x \rightarrow P^1y),$$

whose variables are restricted to terms of $\mathfrak{L}$, is materially true. It follows that the logical status of (12) depends on such mundane facts as the number of individuals in the world. Had the actual world contained only one individual, (12) would have come out logically valid. We will call the general principle underlying this fallacy "the reduction principle" (after Etchemendy). Assuming the notion of logical consequence for $\mathfrak{L}$ is defined in a metalanguage $\mathfrak{L}^* \supseteq \mathfrak{L}$, we can formulate the reduction principle as follows: "If a universally quantified sentence is *true*, then all of its instances are *logically true*."[16] If σ is a sentence of $\mathfrak{L}$ with no non-fixed terms (e.g., $(\exists x)(\exists y)x \neq y$), the reduction principle says that σ is logically true if σ is materially true.

The three fallacies of material consequence are not equal. The first fallacy is caused by an oversimplification of the linguistic test for logical consequence, and it would not be surprising if, with a little technical ingenuity, we were able to obviate it. The second and third fallacies, on the other hand, are inherent in the linguistic conception. It is integral to this conception that the definition of "logical consequence" is couched in a simple, non-modal metalanguage, a language whose notion of truth is strictly material. (This is what sets the linguistic conception apart as an attractive alternative to the metaphysical conception!) But this simplicity means that certain important types of conceptual resources, namely, those required for constructing nonmaterial notions, are missing. Linguistic semantics has neither the means for distinguishing between logical and nonlogical terms, nor the means for reducing logical consequence to a precise yet strong enough notion that would account for its necessity.[17]

Interpretational semantics. Interpretational semantics is generated from substitutional semantics in an attempt to block the first fallacy. It includes a new technical device, the interpretational model, but otherwise it is based on the same principles as its substitutional predecessor. In particular, (i) its notion of logical consequence is relative to arbitrary selections of fixed terms, and (ii) logical consequences are reduced to material generalizations. This means, of course, that interpretational semantics falls prey to the second and third fallacies, and as such it constitutes an unviable alternative to substitutional semantics. But if interpretational semantics is not a viable option, why dwell on it at all? Our interest in interpretational semantics arises from the fact that its technical apparatus is, in certain respects, similar to

[16]Etchemendy [9, p. 98]. Here I apply a rule intended for interpretational semantics to substitutional semantics. This application requires various clarifications that I will not specify in detail here. The idea should be clear: we are referring to substitutional universal quantifications of $\mathfrak{L}^*$ related to sentences of $\mathfrak{L}$ in the way (14) is related to (13); "all instances" refers to instances in $\mathfrak{L}$.

[17]The substitutional definition of logical consequence in Quine [39] avoids the above fallacies by introducing various provisos: (i) the substitution classes include formulas and not just primitive terms, (ii) the fixed terms are restricted to the standard ones minus identity, (iii) only individual variables are permitted, (iv) the (object) language is required to be rich enough for elementary number theory. From the point of view of the present investigation these provisos are ad hoc. For additional problems with the substitutional definition of logical consequence as motivated by nominalism see Putnam [36].

that of mathematical logic, and this might lead one to confuse it with the standard semantics. Based on this confusion one would naturally claim that the standard account of logical consequence "does not capture, or even come close to capturing, any pretheoretic conception of the logical properties" (see above). My goal in discussing interpretational semantics is to remove this confusion.

A natural way to overcome the relativity of substitutional consequences to richness of the lexicon is to create maximal substitution classes. Such classes will contain names for all existent individuals, properties of existent individuals, relations of existent individuals, etc.[18] We arrive at interpretational semantics by introducing a certain technical device – the interpretational model – for obtaining such classes. Given an extensional language, $\mathfrak{L}$, and a selection of primitive terms of $\mathfrak{L}$ to be held "fixed," an interpretational semantics for $\mathfrak{L}$ is a pair, $\langle \mathfrak{L}', \mathfrak{M} \rangle$, where $\mathfrak{L}'$ is a schematic language[19] and $\mathfrak{M}$ is an apparatus of interpretational models for $\mathfrak{L}'$. $\mathfrak{L}'$ is generated from $\mathfrak{L}$ by replacing all the non-fixed primitive terms of $\mathfrak{L}$ by schematic symbols based on syntactic categories: say, proper names are replaced by lowercase Roman letters, a–u, with or without finite ordinal subscripts, n-place predicates of individuals, $n \geq 1$, are replaced by upper case Roman letters, A^n–Z^n, with or without subscripts, etc. Variables play their usual auxiliary role. An interpretational model for $\mathfrak{L}'$ consists of a universe, A, identified with the set of all actual individuals, and an assignment of values to all the non-fixed primitive terms of $\mathfrak{L}$ within A: an individual constant is assigned a member of A, an n-place predicate of individuals is assigned a subset of A^n, etc. A represents a maximal substitution class for the individual constants of the original language, $\mathfrak{L}$; the power set of A^n, $n \geq 1 - \mathscr{P}(A^n)$ – represents a maximal substitution class for all the n-place primitive predicates of individuals of $\mathfrak{L}$; etc. A model for $\mathfrak{L}'$ represents a universal replacement of the primitive non-fixed terms of $\mathfrak{L}$ by primitive terms drawn from a maximal substitution class (in accordance with the rules governing such a replacement: uniformity, respect of syntactic categories, etc.). Instead of saying that "Tarski is a logician" is false *under substitution* of "Sienkiewicz" for "Tarski," we now say that "L^1t" is false *under an interpretation* of "Tarski" as a name of Sienkiewicz or, more technically, that "L^1t" is false *in a model* that assigns Sienkiewicz to t and the set of all actual logicians to L^1. An interpretational model for $\mathfrak{L}'$ is, thus, an interpretation of $\mathfrak{L}$ within a maximal extension $\mathfrak{L}'$ of $\mathfrak{L}$. Contrasting interpretational and representational semantics we can say that representational models represent possible worlds while interpretational models represent possible maximal languages (of the actual world). We define truth in an interpretational model à la Tarski in terms of satisfaction, saying that "L^1t" is true in a model $\mathfrak{A}$ iff the object assigned to t in $\mathfrak{A}$, $t^{\mathfrak{A}}$, *satisfies* the predicate L^1t in $\mathfrak{A}$, or iff $t^{\mathfrak{A}} \in {L^1}^{\mathfrak{A}}$, etc. The interpretational definition of logical consequence (of $\mathfrak{L}'$) is, wordwise, the same as in mathematical logic: σ is a logical consequence of Γ iff every model of Γ is a model of σ (where the notions of logical consequence and model are relativized to $\mathfrak{L}'$). By construction,

[18]Because the linguistic test for logical consequence is based on a material notion of truth, only existent individuals need to be taken into account. If $\mathfrak{L}$ contains names of nonexistent objects, we can reformulate the sentences in which these names occur using predicates.

[19]The language could also be construed as fully interpreted, but since my goal is to show that there are essential differences between interpretational semantics and standard logical semantics I prefer to eliminate the nonessential differences.

interpretational semantics avoids the first fallacy: if there is a language, $\mathfrak{L}^+ \supseteq \mathfrak{L}$, such that some substitution of primitive terms of $\mathfrak{L}^+$ for t, P^1 and L^1 renders "P^1t" materially true and "L^1t" materially false, then there is an interpretational model that represents this substitution.

Interpretational semantics avoids the first fallacy of material consequence but, as I have indicated above, the introduction of maximal substitution classes (in the form of models) does not eliminate the second and third fallacies. The crux of the matter here, as in the earlier case, is the material nature of its meta-theory and the consequent dearth of conceptual resources. Interpretational semantics, just like substitutional semantics, is not equipped for drawing the nonmaterial distinction between logical and non-logical terms. And in interpretational semantics, just as in substitutional semantics, the logical status of inferences like (12) is reduced to the material truth of generalizations like (the objectual analog of) (14).

Is standard logical semantics interpretational? This question is partly ambiguous, since the notion of interpretational semantics involves not just a technical apparatus, but also a theory of what this technical apparatus represents, and while it is clear what the technical apparatus of standard logical semantics is, there is no consensus about its philosophical foundation. (Indeed, such a consensus would be irrelevant anyhow, for the existence of an adequate philosophical foundation for this semantics is at trial here.) However, if we can show that the technical apparatus of standard logical semantics is significantly different from that of interpretational semantics – if, in particular, we can point to differences that (prima facie) preclude the second and third fallacies, we will have provided a (prima facie) ground for a negative answer to this question.

There are two glaring technical differences between the semantic-syntactic system of standard mathematical logic and interpretational semantics: (A) In mathematical logic the selection of logical terms is fixed and definite, while in interpretational semantics the selection of logical terms is variable and arbitrary. (B) In interpretational semantics all models for a given language share the same, predetermined, universe (the "actual" universe, whatever this turns out to be). In mathematical logic, on the other hand, models for a given language vary both with respect to the size of their universes and with respect to the identity of their members (exhibiting the whole gamut of countable and uncountable universes of abstract, physical and fictional objects). These differences have (prima facie) direct implications for the material consequence fallacies. The restriction on logical terms prima facie blocks the second fallacy, while the variations in models constitute a prima facie safeguard against the third fallacy.

Let me dwell on the second point. The difference between the interpretational and the standard notion of *model* is related to differences between the respective meta-theories. Interpretational semantics is defined within a material meta-theory (a theory of "the actual world"), hence its notion of universe is confined to collections of "actual" individuals. The standard semantics of mathematical logic, on the other hand, is defined within a metalanguage that contains a full-fledged abstract set theory, and this theory gives rise to a notion of universe as a well-defined collection (i.e., a set) of any objects not ruled out by the axioms, including the elements of the set theoretic hierarchy. This difference has a radical effect on the respective reduction principles. While interpretational semantics reduces logical inferences to material truths, standard mathematical semantics reduces logical inferences to set theoretical, i.e., mathematical, truths, and mathematical truths, unlike material truths, are intuitively formal and necessary. Thus, whereas interpretational semantics reduces (12) to a sentence whose quantifiers are restricted to actual individuals and collections of actual individuals, the standard semantics reduces (12) to a set theoretical statement, namely

$$(15) \qquad (\forall s_1)(\forall s_2)(\forall x)(\forall y)\{[s_2 \subseteq s_1 \& x \in s_1 \& y \in s_1] \rightarrow [x \in s_2 \rightarrow y \in s_2]\}$$

(where s_1, s_2 range over sets; x, y – over urelements and sets), whose quantifiers are not restricted in this way. The latter, unlike the former, is unencumbered by any material restrictions on the actual universe (e.g., its size).

The fact that the interpretational account implicates standard semantics in fallacies, together with the existence of serious discrepancies between the interpretational account and the data (discrepancies which suggest that the standard system has internal safeguards against the fallacies), creates a strong presumption against the interpretational account. Of course, it is possible to explain away the differences between the interpretational account and the data (we can always nullify gaps between theory and data by making sufficiently radical adjustments).[20] The relevant question is, however, not whether we *can* construe modern semantics interpretationally but, given the price, whether (within reason) we *have* to.

In this section I have examined two conceptions of logic and found them inadequate. The metaphysical conception inflates the intuitive notion of logical consequence by identifying it with that of necessary consequence, making the construction of an adequate logical semantics contingent upon the development of an adequate background metaphysics. The linguistic conception deflates the notion of logical consequence by forcing its reduction to material consequence, thereby turning an intuitively useful and powerful notion into a weak and useless one. In the next section I will propose a conception of logical semantics that, I believe, vindicates Tarski's conjecture. The general principles underlying this conception are based, in large part, on **The Bounds of Logic: A Generalized Viewpoint** [47].

§3. In justifying a theoretical definition of an intuitive concept it is natural to proceed in two steps: (a) justify the general structure of the definition, (b) justify particular applications. In the first step we ask: Is it possible, in principle, to construct a system in which the theoretical definition gives intuitively correct results? What are the general principles underlying such a system? In the second step we inquire whether a particular system satisfies these principles. In the case of Tarski's definition we cannot draw a complete separation between the two since without something like the set-theoretic methods of standard semantics it is almost impossible, practically speaking, to articulate the general principles. My account, however, concerns the general architecture of logical semantics rather than the special design of this or that application, and the intended reading of the set-theoretical terminology is, therefore, as general and as non-committal as (plausibly) possible.

---

[20]This is what Etchemendy seems to be trying to do, though not systematically and partly indirectly. See [9, Chs. 3, 4, 8, 10 and 11].

The thesis I will argue for is relatively strong: Taking Tarski's intuitive conditions as determining the scope as well as the limits of logical semantics, I will develop a conception of logic under which, I will claim, Tarski's definition yields *all* and *only* intuitively logical consequences (relative to a given extensional language). Certain technical developments in logic play an important role in my analysis, in particular, Mostowski's generalization of the standard quantifiers and Lindström's extension of Mostowski's generalization. These generalizations throw new light on the question of logical terms, allowing us (for the first time) to investigate the nature of logical predicates and quantifiers in the same precise and informative manner as in the case of the logical connectives. While Tarski's early ideas on logic and semantics constitute my starting point, my interest in Tarski is not scholarly and I will feel free to diverge from, reshape and extend his ideas in ways that will contribute to the present enterprise.

In broad, and rather crude, lines, my conception can be described as follows: The intuitive notion of logical consequence is that of *necessary and formal* consequence. The key to understanding logical consequence is the formality condition, which allows us to distinguish between the general notion of necessary consequence and the specific notion of logical (= *formally necessary*) consequence. The formal nature of logical consequence is reflected in (i) the choice of logical terms, (ii) the construction of models. Logical terms are formal in the sense of denoting properties and relations that are, roughly, intuitively structural or mathematical. Technically, logical terms do not distinguish between isomorphic arguments, or, more precisely, logical terms are invariant under isomorphic structures (logical terms are those terms whose evaluation commutes with all isomorphisms of domains). Since the concept of logical consequence only takes into account the formal component in the meaning of sentences, the question of identifying logical consequences of a given interpreted language $\mathfrak{L}$ (possibly a segment of "natural" language) reduces to that of identifying the logical consequences of a related, semi-schematic, language $\mathfrak{L}'$ in which only the logical terms are fully interpreted. Models for $\mathfrak{L}'$ represent formally possible structures of objects vis-a-vis $\mathfrak{L}$. In particular, the spectrum of possibilities with respect to the formal features (structures) of objects "detected" by the logical terms of $\mathfrak{L}$ is represented by the apparatus of models for $\mathfrak{L}'$. By construction, consequences preserving truth in all models (for $\mathfrak{L}'$) hold in all formally possible structures of objects (relative to $\mathfrak{L}$), hence are formally necessary (relative to it). Since formal necessity is a particular case of both necessity and of formality, the two intuitive conditions are satisfied. The criterion for logical terms determines a family of logical systems satisfying Tarski's conjecture. We can view this family as a "universal logic": every Tarskian consequence of any system of universal logic is intuitively formal and necessary, and every intuitively formal and necessary consequence (of any "reasonable" language) is identified by some system of universal logic. The notion of *first-order* universal logic is especially fruitful and it partially coincides with "abstract logic" or "model-theoretic logics." Finally, falling within universal logic, standard mathematical logic satisfies Tarski's conjecture.

I will now turn to a detailed presentation of this conception, beginning with Tarski's notion of *semantics*. Tarski arrived at the idea of a semantic definition of "logical consequence" after rejecting the proof-theoretic and substitutional definitions. The proof-theoretic definition, according to Tarski, is too narrow. Consider the following inference (where statements about numbers are high-order, as in the simple theory of types): 0 possesses the property P, 1 possesses the property P, $\ldots$, n possesses the property P, $\ldots$; therefore, all natural numbers possess the property P. Intuitively, this inference is logically valid (formally necessary), but the standard proof-theoretic method fails to recognize its validity. We may contemplate the addition of new rules of inference, but it follows from Gödel's incompleteness theorem that no reasonable proof method (finite rules, etc.) will detect all intuitively logical consequences.[21] The substitutional definition, Tarski noted (based on what I have called the first fallacy of material consequence), reduces the notion of logical consequence to that of material consequence. Having rejected the existent "syntactic" methods for defining logical consequence, Tarski turned to *semantics*, a discipline whose principles were first precisely formulated in his 1933 paper [52].

Semantic reduction. Today, we often view any definition in terms of truth, no matter how truth is defined or what other notions are involved, as a semantic definition. In talking of "linguistic (substitutional, interpretational) semantics" in Section 2, I implicitly accepted this usage. For Tarski, however, the semantic method has to do with a particular way of analyzing concepts, namely, as expressing relations between language and objects: "We shall understand by semantics the totality of considerations concerning those concepts which, roughly speaking, express certain connexions between the expressions of a language and the objects and states of affairs referred to by these expression." [54, p. 401] Henceforth I will restrict myself to this usage. Now, some semantic predicates relate expressions to objects directly: "reference" and "satisfaction" fall under this category. "Truth" and "logical consequence," however, are essentially linguistic predicates, truth being a property of sentences and logical consequence a relation between a sentence and a set of sentences. In what sense are they semantic? One answer to this question turns on definability: truth and logical consequence are definable in terms of (directly) semantic properties, namely, reference and satisfaction, and in this sense they are (indirectly) semantic. But this answer inverts the order of explanation. Truth and logical consequence are definable in terms of reference and satisfaction because they have to do with language in its relation to objects. On my interpretation, truth and logical consequence hold of a given linguistic entity (pair of linguistic entities) in their domain due to (i) certain relations in which the expressions involved stand to certain objects, and (ii) certain facts about these objects."[22] Assuming (i), we can

[21]The fact that Tarski chose a higher-order inference to demonstrate the intuitive inadequacy of the proof-theoretic definition is significant. Etchemendy [8] mistakes this inference for a first-order inference and uses this "fact" to support the claim that Tarski's 1936 notion of logical consequence was different from the common notion. Briefly, Etchemendy argues as follows: The completeness theorem shows that the standard semantic definition of logical consequence does not establish the logical validity of first-order inferences that, like the arithmetic inference Tarski referred to, fail the proof-theoretic test; therefore, Tarski did not have the standard notion of logical consequence (/logical term) in mind. [8, pp. 65, 70–73] It is clear, however, from Tarski's reference [55, p. 410, fn. 1] that the inference he presents is a higher-order inference (see Tarski [53, pp. 279,288]), and in higher-order logic the semantic notion of logical consequence is broader than the proof-theoretic notion. While in Sher [46, pp. 342–3] and [47, pp. 38–9], I, too, read Tarski's inference as a first-order inference, I showed Etchemendy's conclusion could be avoided. My discussion of logical terms below explains why the inference had to be higher-order (why its first-order version is *not* intuitively formally valid).

[22]My use of "object" here is general: individuals are objects, and properties (sets), relations and functions of objects are objects. The main contrast is with "linguistic expression." Of course, linguistic

describe the semantic method as based on the idea that certain linguistic properties – properties of linguistic expressions – are *grounded* in nonlinguistic properties – properties of objects related to these expressions in certain specified ways. In the case of logical consequence, the linguistic relation $\models$ between Γ and σ is grounded in a certain nonlinguistic relation $\mathscr{R}$ between the properties attributed to objects by Γ and those attributed to them by σ (between the configuration of objects described in Γ and that described in σ). What the semantic definition sets out to accomplish is reduction of the former to the latter, and to do so it exploits concepts directly connecting language to objects, like reference and satisfaction.

The nature of the semantic reduction may be obscured by the fact that under certain descriptions the reduction appears to be purely linguistic. Thus, we can describe the reduction of $\models$ to $\mathscr{R}$ by means of a meta-theoretic schema, "$\Gamma \models \sigma$ iff GRS," where "$\Gamma \models \sigma$" stands for a linguistic statement and "GRS" for an objectual statement, and where there is an appropriate reference relation between the object-language expressions referred to in the former and the objects in the world referred to in the latter. Then, if we concentrate on the *mechanics* of writing biconditionals based on this schema (in the simplest case, on the left we write metalinguistic names of object language sentences, on the right – object language sentences or their metalinguistic equivalents), we may come to think of the reduction as based on a certain relation between the metalanguage and the object language. "Logical consequence," we will then say, expresses a "connexion" between expressions of the two languages. But semantic concepts express a "connexion" between language and the *world* ("expressions of a language" and "objects and states of affairs referred to by these expressions"), hence what is philosophically significant about the biconditionals is the fact that one condition is linguistic while the other is objectual. The reduction consists in a passage from a statement about language to a statement about the world: from a statement about a relation between expressions to a statement about a relation between objects. In sum, the semantic reduction grounds a linguistic relation in an objectual one.[23] What kind of objectual relation grounds logical consequence?

Consequence of type χ. There are many kinds of consequence: legal consequence, arithmetic consequence, and so on. Different kinds of consequence are grounded in different relations between objects. Legal consequences are grounded in legal relations, arithmetic consequences are grounded in arithmetic relations, etc. Material consequences, in general, are grounded in material relations, and necessary consequences in necessary relations. To understand what kind of relations ground consequences, let us begin with a non-logical type of consequence, say biological consequence.

Consider the following inferences:

(16) Clinton is a human male; therefore, Clinton evolved from a zygote with one X chromosome.

---

expressions can also play the role of objects, but objects are always distinguished from the expressions which denote them.

[23]I am developing an account of *truth* based on this analysis in *On the Possibility of a Substantive Theory of Truth* [50]. For an abstract of an earlier version see Sher [48].

(17) Clinton is a human male; therefore, Clinton has a progeny.

(18) Clinton is a US President; therefore,
Clinton is a US Commander in Chief.

(19) Clinton is a US President; therefore, Clinton is an Arkansan.

(16) is a biologically valid inference: it is based on a universal biological relation between an individual being a human male and its having a certain chromosomal makeup. (17) is an invalid biological inference: it is based on a contingent biological relation between sex and progeny. (18) and (19) are not biological inferences at all, and from the point of view of biology they are equally invalid. Two factors determine whether a given consequence is a genuine biological consequence: (a) the features of objects taken into account are biological/non-biological; (b) the grounding relation is/is-not biologically universal. In accordance with (a) and (b), a system for determining genuine biological consequences will consist of a language with distinguished biological terms (terms denoting biological features of objects) and a device for distinguishing relations that are biologically universal from those that are not. The idea of a system of models representing all biologically possible structures of objects naturally suggests itself. I will not go into the details of such a system, but the point is that genuine biological consequences are grounded in relations that hold in all biologically possible structures of objects, i.e., in all biological models. These relations hold due to certain biological laws, in particular, laws governing the biological properties denoted by the distinguished terms (in our example: "human male," "X chromosome," "progeny"). Biological models obey these laws, and the assignment of extensions to the distinguished biological terms is based on them (e.g., in all biological models the extension of "x is a human male" is included in that of "x evolved from a zygote with one X chromosome"). In this sense the distinguished terms are said to be fixed. Biological consequences, however, do not depend on the objects and properties denoted by the non-distinguished terms (in our example, "Clinton," "US president," etc.), therefore the extensions of these terms in various models will vary arbitrarily (subject to some constraints of object-expression parity). By construction, genuine biological consequences relative to the distinguished biological terms – biological consequences based on universal biological laws governing the properties, relations and functions denoted by these terms – are truth preserving in all models. To construct a semantic biological system we use a meta-theory that (i) tells us what are all the biologically possible structures of objects, and (ii) determines the extension of the distinguished terms in each biologically possible structure in accordance with the laws of biology. Generally, to construct a semantics for consequences of type χ we (i) select terms of type χ (terms denoting features of objects of type χ) as distinguished terms of the system, and (ii) build an apparatus of models that represent all χ-ally possible structures of objects. In terms of semantic reduction, we reduce "σ is an χ-al consequence of Γ" to "$f_\chi(\Gamma)\mathscr{R}_\chi f_\chi(\sigma)$," where $\mathscr{R}_\chi$ stands for a relation of χ-al (e.g., biological) generality and f_χ is a function that extracts the χ-al (in our example, biological) contents of Γ and σ.

Logical consequence. What distinguishes logical from biological consequences? Clearly, logical consequences are not based on biological features of objects and the space of logical possibilities is different from the space of biological possibilities. What features of objects do logical consequences take into account? What is the space of logical possibilities? We are called upon to identify the property χ such that (a) logical consequences are based on features of objects of type χ, and (b) logical consequences preserve truth in all χ-ally possible structures.

To find χ we search for the most basic intuitive characteristics of distinctly logical consequences. Following Tarski (see citation on the first page of the present paper) I will identify the two central features of logical consequence as *necessity* and *formality*.[24] Necessity, we have seen, is by itself a problematic notion, but formality can be viewed as a modifier of necessity: not all necessary consequences are logical, only *formal*-and-necessary (or *formally* necessary) consequences are. The key to understanding logical consequence is, thus, *formality*.

What is the intuitive notion of formality? It is common to view formality as a purely syntactic notion, and from this point of view the formality of logical consequence is a syntactic, rather than a semantic, feature. Tarski may have unwittingly encouraged this view of formality: logical consequence is formal, Tarski said, in being "uniquely determined by the *form of the sentences* between which it holds" [55, p. 414, my emphasis], and what is sentential form if not a syntactic notion? The notion of sentential form, however, is not an absolute notion. Russell, for example, distinguished between logical and grammatical form: "I met Jones" and "I met a man" have the same (traditional) grammatical form but not the same logical form; [42, p. 168] "I am tired and hungry" and "I am tired and I am hungry" have the same logical form but not the same grammatical form. Sentential form is partly a matter of what a given theory is trying to accomplish, and from the point of view of the semantic definition, syntax is driven by semantic considerations. Let me explain. Not all grammatical distinctions are relevant to a given notion of consequence. For example, the distinction between active and passive voice is irrelevant to biological consequences and, as a result, is not included in the syntax of a theory of biological consequence. The central syntactic notion of a given theory of consequence is that of "distinguished (fixed) term," and this notion is determined based on semantic considerations, namely: what kind of features of objects the given notion of consequence takes into account.

Now, if we already know what the distinguished terms of a given theory of consequence are – what its underlying notion of sentential form is – we can characterize its consequence relation based on this form. In particular, if we assume the standard notion of logical term, we can reconstruct the intuitive notion of logical consequence that (in an idealized sense) led to this choice of distinguished terms. For example, the standard logical terms do not distinguish empirical features of objects, hence logical consequence is not empirical; the standard terms do not distinguish the identity of individuals in a universe of discourse, hence logical consequence does not distinguish it either. And these, indeed, are the very features that Tarski

[24] I will not be able to engage in a historical investigation of Tarski's text here. My interpretation is offered as a way of reading the text that allows us to make good sense of the semantic definition. If the reader's interpretation of Tarski is incompatible with mine, he/she may attribute the proposed analysis to me.

attributed to logical consequence due to its formal nature: "since we are concerned here with the concept of logical, i.e., *formal*, consequence, and hence with a relation which is uniquely determined by the form of the sentences between which it holds [standard logical form], this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence σ or the sentences of the class Γ refer. The consequence relation cannot be affected by replacing the designations of the objects referred to in these sentences by the designations of any other objects." [55, pp. 414–5][25]

However, while taking the logical terms for granted simplified Tarski's task, it also weakened his theory. (See discussion of Tarski's paper in Section 1.) An informative answer to the question of logical consequence requires an informative solution to the problem of logical terms, and for that reason I will not assume a notion of logical term at the outset. Rather than asking: What kind of consequence do the standard logical terms give rise to? – I will pose the question of logical consequence in a more complex form: What kind of consequence, based on what kind of terms, is intuitively logical?

Consider two paradigmatic examples of a logically valid inference.

(20) Everything is predetermined; everything is known in advance; therefore, everything is predetermined and known in advance.

(21) Some Gods are descended from Zeus; therefore, some Gods are descended from Zeus or descended from Hades.

What features of objects (individuals, properties, relations) do these inferences take into account? Clearly not the metaphysical property of being predetermined or the epistemological property of being known in advance. Likewise, neither the identity of Zeus or Hades nor the property of being a God or the relation of being a descendant of play any role in the logical validity of these inferences. The properties and relations that do play a role are those denoted by "or," "and," "some" and "every," and for that reason "or," "and," "some" and "every" play the role of logical terms in these inferences. But what property is denoted by "some"? There are those who believe that no property is denoted by "some": "some" is a syncategorematic term, and syncategorematic terms do not denote. The view of logical terms as syncategorematic expressions may agree with the conventional approach to logic but is incompatible with the semantic approach. Semantics construes the relation of logical consequence as essentially involving objects, but if logical terms are purely conventional, how can they serve as a basis for an objectual relation? The conventional approach to logical terms is, however, not the only approach known to us. Frege analyzed the standard logical quantifiers as standing for "objectual" properties of concepts: properties having to do with the size of their extensions. ("Existence is a property of concepts"; "Affirmation of existence is in fact nothing but denial of the number nought." [11, p. 65]) The objectual tradition was revived (in a somewhat altered form) by Mostowski in

[25](a) See fn. 1 above. (b) The claim that relations definable by means of the standard logical terms (and variables) alone do not "distinguish the identity of individuals in the universe of discourse" was established in Lindenbaum and Tarski [23].

the mid 50's, and has since evolved to a fruitful and influential trend within logic and linguistics. (See Mostowski [31] and Lindstrom [24]; for developments in logic see Barwise and Feferman [2] and the bibliography there; for developments in linguistics see Barwise and Cooper [1], Higginbotham and May [16], Keenan and Stavi [18], van Benthem [58, 59], Gardenfors [13], May [27], Westerståhl [61], etc.) Following this tradition, I will view logical terms as genuinely denoting terms: "Some" denotes the second-level property of being nonempty, "every" denotes the second-level property of being universal, i.e., having an empty complement (in a given universe), "and" (in contexts like those above) stands for the operation of intersection, "or" stands for union.[26]

Now, two common characteristics of the properties (operations) denoted by these terms are *formality* and *generality*. These terms denote properties that are formal and general roughly in the sense of being structural or mathematical and applying to objects (extensions of predicates) in general. Based on this observation, we can explain the logical validity of (20) and (21) as grounded in certain general relations between formal features of objects. (20) is based on the fact that whenever two classes of objects are universal (in a given domain), their intersection is universal (in that domain), (21) is based on the fact that whenever a class of objects is nonempty, its union with a second class of objects is nonempty. These relations intuitively satisfy the pretheoretic requirements of *necessity* and *formality*: (i) they take into account only *formal* features of objects, and (ii) they are based on laws (governing these formal features) that hold in *all formally possible* (hence, possible) structures of objects. (Intuitively, all possible structures of objects are formally possible.) We can say that the relation of logical consequence is grounded in certain general laws of formal structure: "The intersection of universal classes is universal," "The union of a nonempty class with another (possibly identical) class is nonempty," etc. It is not a general law of formal structure that the intersection of nonempty classes is nonempty, hence

(22) Something is predetermined; something is accidental;
therefore, something is predetermined and accidental

is not logically valid. It is a general law of formal structure that class inclusion is transitive, hence

(23) All descendants of Zeus are descendants of Cronus; all
descendants of Cronus are descendants of Gaea; therefore,
all descendants of Zeus are descendants of Gaea

is logically valid. And similarly for other inferences. It is a well-known fact that in standard logic valid inferences are reducible to certain formal (set-theoretical) truths. On the present analysis this is not an accidental by-product of our choice of a medium for formulating logic: biological inferences are reducible to biological truths and logical inferences are (by their nature) reducible to formal truths. Logic, on the present conception, takes certain general laws of formal structure and, using the machinery of logical terms, turns them into general laws of reasoning, applicable

[26]While it is essential for my account that logical terms denote properties, I do not want to commit my account to a particular theory of properties.

in any field of discourse. The fact that biological, physical, psychological, historical, ... structures obey the general laws of formal structure explains the generality ("topic neutrality") of logic: Some reference to formal structure (to complements and unions of properties, identity of individuals, non-emptiness of extensions, etc.) is interwoven in all discourse, and therefore logic (the logic of negation and disjunction, identity, existential quantification, etc.) is universally applicable."[27]

Our analysis has led to the following solution to the problem of logical consequence. The characteristic property of logical consequence (the logical property χ) is *formality*: logical consequences take into account formal features of objects; logical consequences preserve truth in all formally possible structures of objects. Speaking in terms of semantic reduction we can say that Tarski's definition reduces "σ is a logical consequence of Γ" to "$f(\Gamma)\mathscr{R}f(\sigma)$," where f is a function (based on the logical terms) that "picks" the formal content or skeleton of what the sentence σ and the "theory" Γ say about objects, and $\mathscr{R}$ stands for the relation of formal generality, roughly "whenever," ranging over all formally possible structures of objects.

A system for identifying logical consequences. To construct a semantic apparatus for identifying logical consequences based on this reduction, the "Tarskian" logician builds a system that consists of a language and a machinery of models. The primitive vocabulary is divided into logical terms, terms representing those formal features of objects that the system takes into account (for example, in standard mathematical logic, being nonempty), and non-logical terms, terms representing nonformal features of objects and formal features not taken into account by the system (for example, in standard first-order logic, being green and being finite). The model-theoretic machinery consists of an apparatus of models and a definition of the logical terms. Intuitively, a (logical) model represents a formally possible structure of objects relative to the primitive terms of a given language. All and only features of objects "detectable" by the logical terms need to be represented by models: identities and differences of individuals, complements of extensions (which requires a specification of a universe), certain cardinality features, etc. Since all the relevant features are extensional, objects are represented in models extensionally. Models are subject to formal constraints (e.g., a set and its complement have no common elements) but not to metaphysical, physical, behavioral, ... constraints. The non-logical terms are *strongly variable*: any formally possible denotation in accordance with their formal "skeleton" (i.e., their being individual terms, n-place relation of individuals, etc.) is represented in some model. The logical terms, on the other hand, denote *fixed* formal properties of objects, and their denotations are subject to the laws governing these properties. These terms are fixed not in the sense that they denote the same entity in each model (the denotation of the universal quantifier in a model with 10 elements differs from its denotation in a model with 11 elements). Rather, while non-logical terms are defined *within* models, logical terms are defined by fixed functions *over* models. Looking at the system as a

[27] I will not go into epistemological issues concerning the background theory of formal structure here. On the relationship between the development of mathematical and logical theories see Sher [47, pp. 133–4]. One way to challenge a given system of logic, on the present analysis, is to challenge the universal applicability of its underlying notion of formal structure. The challenge of fuzzy logic to "classical" logic is naturally viewed along these lines (fuzzy vs. classical sets).

whole, we may say that the primitive non-logical terms and their denotations in models constitute the *base* of the system; the logical terms and their semantic definitions – its *superstructure*. The two parts of the system are brought together by "superimposing" the logical apparatus on the non-logical base. Syntactically, this is done by the definition of well-formed formulae (wffs), in which the logical terms are distinct formula building operators; semantically, by the definition of "truth in a model" which is based on (i) the logical structure of wffs, (ii) the model-theoretic definitions of the logical-terms. Thus, the definition of truth in a model says that "$\exists x\Phi x$" is true in $\mathfrak{A}$ iff the extension of "Φx" in $\mathfrak{A}$ is not empty, etc. By construction, consequences preserving truth in all models, i.e., consequences satisfying Tarski's definition, are formal and necessary, i.e., logical. I will call a logic constructed in the way just described a *Tarskian logic*.

A background theory of formal structure. Tarskian logic is embedded in a general theory of formal structure. The "totality" of models, the definition of logical terms, the extension of "truth in $\mathfrak{A}$," are all determined based on this theory. In standard logic the background theory is ZFC, but from the point of view of the present conception this is not essential. The general architecture of "Tarskian" semantics requires a background theory of formal structure but it is not tied to ZFC (or to any other specific theory). On the other hand, ZFC (unlike relativity theory, Greek mythology, or even group theory) is the right kind of theory for this purpose.

A criterion for logical terms. Logic, on my account, is a theory of formal reasoning, and it is the job of logical terms to represent the formal properties and relations on which this reasoning is based. Intuitively, the distinguished terms of standard mathematical logic satisfy this requirement, but our analysis indicates that any formal property can serve as a basis for formally necessary, i.e., logical, consequences. Take, for instance, finiteness. Consider the inferences

(24) Earth has exactly one satellite; therefore, Earth has less than ten satellites,

and

(25) Earth has exactly one satellite; therefore,
Earth has only finitely many satellites.

The two inferences are equally formal and necessary and both share the same logical form (in both, both premise and conclusion attribute some cardinality property to a certain set of objects). Yet, from the point of view of standard first-order logic (24) is valid while (25) is not, and this result is directly connected to the fact that "less than ten" can be defined in terms of the logical constants of this logic while "finitely many" cannot. What expressions can play the role of distinguished terms in logic?

Two natural criteria for logical terms are: (a) logical terms are structural, (b) logical terms are mathematical. There is a clear similarity between my view of formal structure and the structuralists' view of structures (e.g., Resnik [41]). Starting with a particular situation, say, the child Danny and his six "action figures," we abstract from the particular objects and properties present and we obtain a *structure*, or in Resnik's terminology, a *pattern*, of seven individuals, one of them distinguished and standing to all the other individuals in a certain relation R. This pattern fits many situations: Henry the Eighth and his six wives, the number six and the natural numbers smaller than it, etc. In each of these structures we have *seven* individuals, in each R is *not empty*, in each R has a *complement*, hence "seven," "some," "not" are structural notions. Abstracting from the number of individuals in the above structure we obtain structures with different numbers of individuals and, accordingly, new structural notions: "three," "ten," "$\aleph_0$," etc. However, from my point of view it is important that there are many kinds of structure, obtained by abstraction from different kinds of features. Thus, biological concepts (e.g., "mammal") are obtained by abstraction, and biological inferences, in general, involve abstraction. (E.g., "Cerberus is a mammal; therefore, Cerberus has a heart" does not take into account the identity of Cerberus). Any theory can be viewed as a description of some structure, but not all structures play a role in defining the notion of *logical* constant. To distinguish the structures and structural features that play a role in logic I talk about *formal* structures and *formal* objects (properties, relations and functions of objects). This leads to our second proposal.

It would be natural to identify "formal" with "mathematical" and, following the structuralists, define the identity conditions of mathematical structures in terms of isomorphism. This suggestion comes close to capturing my intention, but certain clarifications are required. Consider the mathematical statement "two is not equal to three." Standard mathematical logic is ambivalent with respect to this sentence: depending on how we construe number expressions – as individual constants, first-order quantifiers, or predicates of standard higher-order logic – this sentence comes out either logically true or logically indeterminate. Are the numerals formal terms? Are they formal under some but not all construals? Which ones? Why?

In his 1957 paper, *On a Generalization of Quantifiers*, Mostowski proposed a condition on quantifiers that, in my view, captures the intuitive idea of formality: "[Formal] quantifiers should not allow us to distinguish between different elements of [the universe of discourse]." [31, p. 13] Defining a quantifier as a 1-place operator which, given a universe (i.e., nonempty set) A, denotes an A-quantifier – a function from the power set of A, $\mathscr{P}(A)$, to T,F – Mostowski interpreted the above condition as saying that a [formal] A-quantifier is invariant under permutations of $\mathscr{P}(A)$ induced by permutations of A. Mostowski's quantifiers correspond to 1-place second-level predicates whose arguments are 1-place (first-level) predicates. More generally, we can define such a predicate, $\mathcal{P}$, as denoting, in any given universe, A, a set of subsets of A, $\mathcal{P}_A$. $\mathcal{P}$ is a formal predicate iff for any universes A and A', and any subsets B and B' of A and A' respectively: if $\langle A, \langle B \rangle \rangle \cong \langle A', \langle B' \rangle \rangle$, then $B \in \mathcal{P}_A$ iff $B' \in \mathcal{P}_{A'}$. This criterion captures the intuitive idea that a formal term distinguishes only formal features of its arguments: $\mathcal{P}$ is a formal predicate iff $\mathcal{P}$ does not distinguish between isomorphic arguments. The same characterization can be applied to terms in general:

(F) A term is formal iff it is invariant under isomorphic structures [24, 57],

or, more informatively, a term is logical iff its evaluation commutes with all isomorphisms of domains. A term with no arguments is not formal in the sense that it does not distinguish any features of objects, formal or nonformal.[28]

[28]For more extended discussions and conditions on the definition of logical (formal) terms, see Sher [46] and [47, Ch. 3]. A response to a challenge due to McCarthy [28] appears in Sher [47, p. 64]. A

We can now explain the difference between the individual constant "two" and the quantifier "two" by saying that the latter, but not the former, is a formal term. Speaking in terms of objects we can say that formal objects are not just elements of formal structures, they are themselves formal structures. Among the terms satisfying (F) are identity and the standard quantifiers, all n-place cardinality quantifiers ("There are at-least/exactly/at-most δx such that," where δ is any cardinal, "There are finitely/denumerably/non-denumerably many x such that," "most x such that ... , – ," etc.). Other formal terms based on (F) are the second-level well-ordering relation, the second-level membership relation, and so on. The first-level predicates "x is green," "x is a number," "x is a member of y" etc., are not formal. Neither are the second-level predicates "X is a property of Napoleon," "R is a relation between physical objects," and others. To see why the first-level membership relation is not formal, consider the two structures $\langle A, \langle \text{Tarski, Gödel}\rangle\rangle$ and $\langle A', \langle \Phi, \{\Phi\}\rangle\rangle$, where $A = \{\text{Tarski, Gödel}\}$ and $A' = \{\Phi, \{\Phi\}\}$. These structures are isomorphic, yet $\{\Phi, \{\Phi\}\} > \in$"is a member of"$_A$ but $\langle \text{Tarski, Gödel}\rangle \notin$"is a member of"$_A$.

When it comes to sentential connectives, we can regard their formality as based on "not distinguishing the identity of propositions." Intuitively, sentential connectives are formal iff they distinguish only patterns of propositions possessing truth values and nothing else. The interpretation of logical connectives as (denoting) Boolean truth functions reflects just this intuition.

So far I have characterized formal constants in terms of the features of objects they "detect." We can also characterize them in terms of the structures they "generate." Thus, applied to the first argument of a 2-place relation $R \in A^2$, the universal quantifier generates a set that stands in a certain formal relation to R, namely, $\{b \in A : \{a \in A : aRb\}\text{is universal in } A\}$; applied to $B \subseteq A$, the negation operator generates a set that stands in a certain formal relation to B, namely, the complement of B in A, etc. Invariance under isomorphic structures can be used as a formality criterion for all "structure generators."[29]

Universal logic. Given an interpreted language $\mathfrak{L}$ (a natural, scientific or mathematical language), we categorize the primitive vocabulary of $\mathfrak{L}$ in a "Fregean" manner, i.e., (i) the type of a term is determined based on the type of its arguments, (ii) no prior division of terms to logical and non-logical is assumed. We correlate the syntactic typology with a semantic (objectual) typology (relative to a universe A) in the natural way: the syntactic type of an individual constant is correlated with the semantic type of a member of A; the syntactic type of an n-place predicate of

"constructive" definition of logical terms based on (F) is given in Sher [47, Ch. 4]. (An abridged version is presented in Sher [49].) Wilfrid Hodges notes the following concise formulation of the definition of formal terms for logics of an arbitrary order: Let A be a set. The *universe over* A, $\mathscr{V}(A)$, is defined by induction on the ordinals: $\mathscr{V}_0(A) = A$; $\mathscr{V}_{\alpha+1}(A) = V_\alpha(A) \cup \mathscr{P}(\mathscr{V}_\alpha(A))$; $\mathscr{V}_\delta(A) = \cup_{\alpha<\delta}\mathscr{V}_\alpha(A)$, when δ is a limit ordinal. $\mathscr{V}(A) = \cup_{\alpha \text{ ordinal}}\mathscr{V}_\alpha(A)$. A term C is formal iff for any nonempty sets A, B: if $e : A \mapsto B$ is a bijection, then $e^*(C_A) = C_B$, where e^* is the isomorphism of $\mathscr{V}(A)$ onto $\mathscr{V}(B)$ induced by e. (The use of "universe" here is different form that in the text. Below, I use "universe" to refer to a nonempty set under a certain role ("the universe of a structure," or "the universe of a model").) The notion of formal term can be naturally extended to infinitistic languages.

[29](a) From this point of view we interpret the standard logical connectives primarily as (denoting) set operators and only secondarily as (denoting) truth functions. (b) The invariance criterion for structure generators is based on the principle that if an n-place term C is a formal structure generator, then for all A, A': $C^n_{A'}(D'_1, \ldots, D'_n) = D'$ if $C^n_A(D_1, \ldots, D_n) = D$ and $\langle A, \langle D_1, \ldots, D_n, D\rangle\rangle \cong \langle A', \langle D'_1, \ldots, D'_n, D'\rangle\rangle$.

individuals is correlated with the semantic type of a member of $\mathscr{P}(A^n)$, etc. (I will not go into details here.) We define the notion of an *eligible logical term* of $\mathfrak{L}$ based on (F) as follows:[30]

(EL) $\mathscr{C}$ is an eligible logical term of $\mathfrak{L}$ iff $\mathscr{C}$ is either a truth functional connective of $\mathfrak{L}$ or a predicate of $\mathfrak{L}$ satisfying (F).

We select a set, $\mathcal{E}$, of *logical terms* of $\mathfrak{L}$ based on (EL),[31] and we construct a logical system, $\mathscr{L}$, in the way described above. (The set of logical constants of $\mathscr{L}$ represents $\mathcal{E}$; its set of non-logical constants represents the primitive non-logical vocabulary of $\mathfrak{L}$.) We apply Tarski's test to $\mathscr{L}$ and, translating back to $\mathfrak{L}$, we obtain the the set of logical consequences of $\mathfrak{L}$ relative to $\mathcal{E}$.

The criterion (EL), in abstraction from $\mathscr{L}$, leads to the idea of logic as a *family* of systems for identifying formal and necessary, i.e., logical, consequences. The various systems differ in their logical constants, but all share the property of being formal in the sense that (i) their logical predicates satisfy (F), (ii) their logical connectives are truth-functional, and (iii) their apparatus of models represents all formally possible structures of objects relative to the logical (formal) vocabulary. We may think of this family of logical systems as *universal logic*.[32] Universal logic represents a certain totality of logical systems based on the "formal" or "Tarskian" conception of logic. If we restrict ourselves to the family of logics with a non-logical vocabulary of levels 0 and 1 (individual terms and predicates of individuals), and a logical vocabulary of truth-functional connectives plus predicates of levels 1 and 2 (predicates with arguments of levels 0 and/or 1), we arrive at a class of logics "for" languages with a first-level non-logical vocabulary. I will call this class *universal first-order logic*. The notion of universal first-order logic partly coincides with those of *model theoretic* and *abstract* logic. Its fruitfulness is demonstrated by prolific meta-logical results: Lindström's theorems, various completeness and incompleteness results, etc., and numerous mathematical and linguistic applications. (See articles and references in Barwise and Feferman [2] and Gärdenfors [13].) The term "first-order logic" is often reserved for *standard* first-order systems, but from the present perspective any system of universal first-order logic is a first-order logic. (We may view the standard system as "elementary" first-order logic.) The Löwenheim-Skolem theorem points to a certain redundancy in the standard first-order notion of model, but this redundancy disappears when we think of models within the framework of universal first-order logic. Indeed, we can view the Löwenheim-Skolem theorem as an incompleteness theorem for standard first-order logic: certain intuitively formal features of objects, in particular, those having to do with uncountable cardinalities, are not detected by

[30](a) Strictly speaking, there is no need to separate the cases of connectives and predicates. See fns. 28 and 29 above, as well as Lindström [24]. (b) (EL), and the associated notion of "universal logic," are not intended to detract from the value of intensional logics. My analysis, however, does point to a difference between the philosophical principles underlying mathematical logic and those underlying the various intensional logics.

[31]Some predicates of the natural/scientific/mathematical language $\mathfrak{L}$ may not have a determined extension in every universe A, and this could interfere with the application of (EL). There are, however, familiar conventional solutions to this problem.

[32]In Sher [47] I used the term "unrestricted logic," abbreviated as "UL." Several of my students read "UL" as "universal logic," and after thinking the matter over I decided to adopt their reading.

this logic. It takes a broader system (family of systems) of logic to account for the notion of first-order logical consequence – consequence based on formal features of individuals and their properties (relations) – in full generality.

Questions and objections. I will clarify certain aspects of the methodology of my conception by responding to possible questions and objections.

Q: On your account the extension of "logical consequence" depends on truths of some non-logical background theory. But isn't logic prior to (stronger than, more certain than) all other theories? Doesn't your account lead to an unduly weak notion of logic? **A:** This question raises an important methodological issue that I will only be able to touch upon briefly here. The demand that logic should not depend on any non-logical theory comes from a foundationalist approach to logic. My own view of logic is, however, non-foundationalist. This is not just an external preference that I bring into my discussion of logic, but it is an integral part of the Tarskian project. This project starts with certain intuitive characteristics of logical consequence, and turns them into substantive (informative, non-trivial) conditions on an adequate definition of this logical concept. Foundationalism does not permit us to approach the definition of logical concepts in this way. On the foundationalist approach logic provides a foundation for other disciplines, but nothing (other than logic itself) provides a foundation for logic. The foundationalist approach encourages an all-or-nothing attitude towards logic: being an ultimate foundation, logic is viewed either as something that does not require a justification (i.e., we accept logic on faith, based on some infallible intuition, etc.), or as something that cannot be justified (i.e., we accept, or discard, logic as arbitrary, accidental, conventional). Either way, the foundationalist does not ask, and does not explain, why logic is the way it is. The non-foundationalist approach of the present paper, on the other hand, demands an informative account of logic. It requires that we think of logic in critical and constructive terms, and it does not accept a hand waiving response to such basic questions as the question of logical terms. Logic is thought of as a dynamic discipline whose development is shaped by the ingenuity and imagination of its developers, the changing conceptions of its nature and purpose, problems both internal and external (e.g., applicability to, or compatibility with, other disciplines), etc. On a non-foundationalist approach logic is, of course, not *above* our theories of formal structure (or, for that matter, any theory) and the point of constructing a definition of "logical consequence" is to systematize, or develop, a certain interesting and useful notion of consequence; Tarski's notion of logical consequence is not stronger than that of formal-and-necessary, i.e., broadly speaking, mathematical consequence, and the relation between logic and mathematics is that of interrelated disciplines, the advancement of one contributing to the advancement of the other. The fact that the emergence of modern logic coincided with the development of a rigorous set theory is not accidental, on this view: logic requires a powerful background theory of formal structure as much as set theory requires a powerful logical machinery.[33] There are, of course, certain methodological constraints on the relations between logic and mathematics. For example, we cannot reduce logic to mathematics and then reduce mathematics to logic. However, my approach is

[33]See Vaught [60] for another aspect of the relation between modern semantics and set theory.

compatible with a wide variety of philosophies of mathematics (various strands of structuralism, Platonism, constructivism and naturalism.)

Q: Tarski's definition reduces logical consequence to preservation of truth in all models. But how do you know that no accidental feature is common to all models? **A:** The roots of this objection can be traced to Wittgenstein who said: "The mark of a logical proposition is *not* general validity. To be general means no more than to be accidentally valid for all things." [62, 6.1231] While Wittgenstein's objection applies to the interpretational conception of logic, it does not apply to my conception, where logical consequence is reducible not to just any kind of generality, but to a special kind of generality, namely, *formal and necessary* generality.[34] Speaking in terms of models: Suppose there is an accidental property H, of all models for a given language. The notion of model is defined within some background theory, $\mathfrak{T}$, based on its notion of "formal structure." If $\mathfrak{T}$ is an adequate theory of formal structure, then $\mathfrak{T}$ includes the theorem "Some formal structure $\mathfrak{A}$ does not possess the property H" and, in accordance with this theorem, the apparatus of models defined in $\mathfrak{T}$ will include a model representing a formal structure in which H does not hold. Similarly, if H is a necessary feature of formal structures, $\mathfrak{T}$ will require that all formal structures possess the property H, in which case all Tarskian models will possess this property.

Q: The vindication of Tarski's conjecture, according to your analysis, requires the existence of an adequate background theory of formal structure. How do you know that such a theory exists? **A:** In evaluating a conception of logic within a non-foundationalist methodology we have to distinguish the task of justifying its general structure and the task of justifying its background theories. An extreme *holist* might say that we cannot justify one theory within our total system of knowledge without justifying all the others, but the same motivation that led me to reject the foundationalist approach (its being an all-or-nothing approach) leads me to reject extreme holism. My project is not that of justifying logic by justifying the whole of knowledge. The challenge I responded to, in writing this article, is a challenge to the general architecture of logic, and many questions concerning the use of a particular background theory, though interesting and important, do not directly pertain to this challenge. Of course, if someone had shown that the very idea of a general theory of formal structure is deeply flawed, that would have undermined my conception. But no one, to the best of my knowledge, has shown this. On the contrary, the success of axiomatic set theory in overcoming Russell's paradox, the fact that no new paradoxes have been discovered, the relative consistency of the axioms of ZFC, the agreement of modern set theory with the whole of classical mathematics, the existence of several potentially viable alternatives to the set-theoretical approach to formal structure (Hellman's modal structuralism, Chihara's constructible mathematics, etc.), yield support to my presumption that an adequate theory of formal structure is, in principle, possible, and substantial portions of such a theory already exist.[35]

Q: Set theory, and other theories that can serve as a basis for a theory of formal structure, are theories of *existent* formal structures, not theories of *formally possible*

[34] García-Carpintero [12] regards Wittgenstein's objection as underlying Etchemendy's criticism. Wittgenstein himself goes on to say that "the general validity of logic might be called essential" [62, 6.1232]. In my view, logical validity is grounded in a special kind of "essential generality."

[35] See Boolos [4, 5] for issues pertaining to the adequacy of ZF as a background theory.

structures. **A:** The reductive approach to modal notions is based on the idea that imprecise intuitive notions can be reduced to clear yet adequately strong precise notions. The success of the reduction is measured by (i) the simplicity and clarity of the reductant notions and (ii) their ability to support the strong modal notions. Metaphysical semantics fails to satisfy (i), linguistic semantics fails to satisfy (ii). On my conception, formal possibility is reduced to mathematical existence and formal necessity to mathematical generality: "It is formally possible that Φ " to "There exists at least one mathematical (set theoretical) structure $\mathfrak{S}$ such that Φ holds in $\mathfrak{S}$" and "It is formally necessary that Φ " to "For all mathematical (set theoretical) structures $\mathfrak{S}$, Φ holds in $\mathfrak{S}$." While there are many ways to look at mathematics in general and set theory in particular, the relevant question in the present context is whether set theoretical existence and generality are both sufficiently clear and sufficiently strong to account for the intuitive notion of formal possibility. I will not attempt to decide between ZFC and other theories of formal structure here, but my view is that the notion of existence in ZFC satisfies (i) and (ii), hence ZFC is a reasonable candidate for the reduction of logical consequence. In this context, I read the axiom of infinity as saying that an infinite set (structure) of objects is formally possible, I read Cantor's theorem as saying that it is formally necessary that there are more subsets of a given set (or structure) of objects, $\mathfrak{X}$, than members of $\mathfrak{X}$, and so on.[36] Now, we may not be able to axiomatize our idea of formal necessity/possibility with complete generality (Gödel's incompleteness theorem), but I do not regard this as a serious threat to my conception. If our theory of logic is based on the idea that logical consequence is formally necessary, then it is bound to reside within the limits of our ability to understand (and systematize) the intuitive notion of formal necessity. This does not mean that anything goes, but that in the end we have to judge the adequacy of a background theory for logic based on relatively, not absolutely, stringent standards. On a positive note, this means that logic stands to benefit from advances in related disciplines, for example, a solution to the continuum problem.[37]

Q: There is a sense in which the decision on a background theory for logic rests on contingent matters, for example, the question whether the universe is finite or infinite. Intuitively, both finitism and "infinitism" are logically consistent. **A:** First, I would like to state the obvious, namely, only a controversy with regard to the size of *formally possible* structures, not materially existent structures, is relevant to my conception. Second, the controversy over infinite structures stands to predicate logic in the same way that the controversy over bivalence stands to sentential logic. To use bivalence as a background assumption for standard sentential logic is *not* to commit ourselves to the view that many-valued theories of truth are inconsistent. Similarly, to use infinitism as a background assumption of universal logic is not to claim that finitism is inconsistent. (Indeed, we can formulate finitist theories within an infinitistic logic.) But the same need not hold in the opposite direction. If (in

[36] My interpretation of mathematical theorems in the present context is "modalistic," but note the direction of explanation: it is mathematics that explains (systematizes, clarifies) the notion of formal possibility rather than the other way around. For a discussion of modalism see, for example, Parsons [33]. On the direction of explanation see Putnam [35].

[37] For example, if we solved the continuum problem we could apply Tarski's definition within logics containing the pair of quantifiers $(2^{\aleph_0}x)$ and $(\aleph_1 x)$.

the extreme case) the finitist is saying that an infinite structure of objects is formally impossible, then a theory requiring infinite models is, for him/her, unsatisfiable. Someone may feel that formal possibility is not pure enough for logic, that logical concepts should be defined in terms of "purer" notions. But aside from a trivial reduction of logic to logic it is not clear to me what notions would qualify. One of the points I have tried to make in this paper is that the task of accounting for logic does not force us to choose between an impossible problem and an inept solution: between some kind of rarefied, lofty explanation and a purely material explanation.

§4. Did Tarski commit "Tarski's Fallacy"? Has modern logic committed a fallacy? In this paper I have examined Tarski's definition of logical consequence with respect to three conceptions of logic. Etchemendy's claim that Tarski's definition yields a material rather than a logical notion of consequence holds for the interpretational (or, more generally, linguistic) conception, but the view that modern logic is an interpretational theory is not supported by the evidence. Etchemendy seems to derive this view (at least in part) from a simple disjunctive argument: Either modern logic is interpretational or modern logic is representational; modern logic is not representational; therefore, modern logic is interpretational. The dilemma on which this argument rests is, however, inherently problematic. How do we justify the claim that there are just two (viable) interpretations of modern logic? Etchemendy does not bring any explicit argument in support of his dilemma. Implicitly, he may be reasoning as follows: Logical consequence is based either on language or on the world; interpretational semantics is *the* theory that logic is based on language; representational semantics is *the* theory that logic is based on the world; ergo the dilemma. But this reasoning involves an uncritical uniqueness assumption: there is exactly one theory of logic as based on language and exactly one theory of logic as based on the world. Furthermore, the very idea of logic "being based on language" or "being based on the world" is simplistic. There are many aspects of language (the aspect studied by the grammarian, the aspect studied by the stylist, etc.) and many aspects of the world (the aspect studied by the physicist, those studied by the biologist, metaphysician, etc.). The question is not simply whether logic is based on language or the world; the question is what *aspect* of language and what *aspect* of the world (indeed, what aspect of the language-world relation) logic is based on.

Both the representational and the interpretational answers to the question of logic are inadequate. Representational semantics construes logic as based on general metaphysics (metaphysical features of objects), but "logical consequence" is not the same notion as "metaphysical (necessary) consequence." Using a notoriously abstruse and unwieldy notion to define a comparatively clear and transparent one is methodologically unsound. Interpretational semantics is an attempt to derive the notion of logical consequence from the idea of arbitrary distinguished terms by means of a theory of material truth. Interpretational semantics, however, tries to do too much with too little, and not surprisingly its notion of logical consequence collapses to material consequence.

One of the main obstacles to developing an adequate account of logic has been the characterization of "truth in virtue of form." In his introduction to the second edition of **The principles of mathematics**, Russell said: "It seems clear that there must be some way of defining logic otherwise than in relation to a particular logical language. The fundamental characteristic of logic, obviously, is that which is indicated when we say that logical propositions are true in virtue of their form. ... I am unable to give any clear account of what is meant by saying that a proposition is 'true in virtue of its form.' But this phrase, inadequate as it is, points, I think, to the problem which must be solved if an adequate definition of logic is to be found." [43, xii] The formal conception of logical consequence outlined in this paper offers a definite solution to this problem, culminating in a precise and unequivocal criterion for logical constants, namely (EL). Etchemendy [9, Ch. 9] discards the issue of logical constants as a "red herring," but it is important to keep in mind the context in which his claim is made. Etchemendy assumes that modern logic is an interpretational system, and on this assumption the view that a demarcation of logical terms will rescue logic from the fallacies is indeed a "myth": interpretational semantics falls prey to the third as well as the second material consequence fallacy, and a solution to the problem of logical terms will not avert the third fallacy.

My philosophical conception owes much to internal developments in logic. Mostowski's generalization of the logical quantifiers and Lindström's (and Tarski's) extension of Mostowski's generalization created a new environment for meta-logical investigations. By offering a genuine, systematic alternative to the standard selection of logical terms, this environment has made it possible to conduct a general philosophical critique of the notion of logical term and with it, in a sense, the notion of logic. While Mostowski's and Lindström's generalizations have had a profound influence on the development of mathematical logic as well as linguistic semantics, many philosophers still restrict their deliberations to standard first-order logic. My account closes the gap between the philosopher's and the working logician's (and linguist's) notions of logical semantics. I believe that, in addition to profuse mathematical and linguistic results, the new, broader perspective on logic has numerous philosophical ramifications: for ontology (ontological commitment and ontological relativity), for the philosophy of mind ("the logic of thought," mind/brain computation, for the theory of truth, for the theory of meaning, and for the philosophy of mathematics (e.g., logicism). I have examined some of these ramifications in Sher [47, Ch. 6]. See also Shagrir [44, Ch. 2]

Due to limitations of space I have not discussed Tarski's 1936 [55] paper in detail here. I have pointed out that Tarski, as far as we know, never identified the proof referred to in his 1936 paper, and I have concluded that, in the absence of any supporting evidence, the claim that Tarski committed "Tarski's fallacy" is unsubstantiated. There is, however, a widespread (though largely undocumented) view that Tarski "got things wrong" in his 1936 paper. The examination of this view must wait for another occasion.[38] In the present paper I have investigated the modern definition of logical consequence that emerged from Tarski's paper. This definition, I hope I have shown, is coherent and well motivated and, within the bounds of "universal logic," it does captures the intuitive notion.

[38]For one explanation of presumed gaps in Tarski [55] see Hodges [17].

REFERENCES

[1] J. BARWISE and R. COOPER, *Generalized quantifiers and natural language*, ***Linguistics and Philosophy***, vol. 4 (1981), pp. 159–219.

[2] J. BARWISE and S. FEFERMAN, ***Model-theoretic logics***, Springer-Verlag, New York, 1985.

[3] B. BOLZANO, ***Theory of science***, D. Reidel, Dordrecht, 1973.

[4] G. BOOLOS, *On second-order logic*, ***Journal of Philosophy***, vol. 72 (1975), pp. 509–27.

[5] ———, *Nominalist platonism*, ***Philosophical Review***, vol. 94 (1985), pp. 327–44.

[6] C. CHIHARA, ***Constructibility and mathematical existence***, Oxford, 1990.

[7] H. B. ENDERTON, ***A mathematical introduction to logic***, Academic Press, New York, 1972.

[8] J. ETCHEMENDY, *Tarski on truth and logical consequence*, this JOURNAL, vol. 53 (1988), pp. 51–79.

[9] ———, ***The concept of logical consequence***, Harvard, Cambridge, 1990.

[10] G. FORBES, ***The metaphysics of modality***, Clarendon, Oxford, 1985.

[11] G. FREGE, ***The foundations of arithmetic***, Northwestern, Evanston, 1884, translated by J. L. Austin, 1986.

[12] M. S.-M. GARCÍA-CARPINTERO, *The grounds for the model-theoretic account of the logical properties*, ***Notre Dame Journal of Formal Logic***, vol. 34 (1993), pp. 107–31.

[13] P. Gardenfors (editor), ***Generalized quantifiers: Linguistic and logical approaches***, Reidel, 1987.

[14] D. GOTTLIEB, ***Ontological economy***, Oxford, 1980.

[15] G. HELLMAN, ***Mathematics without numbers***, Oxford, 1989.

[16] J. HIGGINBOTHAM and R. MAY, *Questions, quantifiers and crossing*, ***Linguistic Review***, vol. 1 (1981), pp. 41–79.

[17] W. HODGES, *Truth in a structure*, ***Proceedings of Aristotelian Society*** (1986), pp. 135–51.

[18] E. L. KEENAN and J. STAVI, *A semantic characterization of natural language determiners*, ***Linguistics and Philosophy***, vol. 9 (1986), pp. 253–329.

[19] H.J. KEISLER, *Logic with the quantifier 'there exist uncountably many'*, ***Annals of Mathematical Logic***, vol. 1 (1970), pp. 1–93.

[20] G. KREISEL, *Informal rigour and completeness proofs*, ***Problems in the philosophy of mathematics*** (I. Lakatos, editor), North-Holland, Amsterdam, 1969, pp. 138–71.

[21] S. A. KRIPKE, ***Naming and necessity***, Harvard, Cambridge, 1972.

[22] ———, *Is there a problem about substitutional quantification?*, ***Truth and meaning*** (Evans and McDowell, editors), Oxford, 1976, pp. 325–419.

[23] A. LINDENBAUM and A. TARSKI, *On the limitations of the means of expression of deductive theories*, in Tarski [1983], pp. 382–92.

[24] P. LINDSTROM, *First order predicate logic with generalized quantifiers*, ***Theoria***, vol. 32 (1966), pp. 186–95.

[25] R. B. MARCUS, *Interpreting quantification*, ***Inquiry***, vol. 5 (1962), pp. 252–9.

[26] ———, *Quantification and ontology*, ***Noûs***, vol. 6 (1972), pp. 240–50.

[27] R. MAY, *Interpreting logical form*, ***Linguistics and Philosophy***, vol. 12 (1989), pp. 387–435.

[28] T. MCCARTHY, *The idea of a logical constant*, ***Journal of Philosophy***, vol. 78 (1981), pp. 499–523.

[29] V. MCGEE, *Review of the concept of logical consequence*, this JOURNAL, vol. 57 (1992), pp. 254–5.

[30] ———, *Two problems with Tarski's theory of consequence*, ***Proceedings of The Aristotelian Society***, vol. 92 (1992), pp. 273–92.

[31] A. MOSTOWSKI, *On a generalization of quantifiers*, ***Fundamenta Mathematicae***, vol. 44 (1957), pp. 12–36.

[32] C. PARSONS, *A plea for substitutional quantification*, 1971, ***Mathematics in philosophy***, Cornell, 1983.

[33] ———, *The structuralist view of mathematical objects*, ***Synthese***, vol. 84 (1990), pp. 303–46.

[34] C. PEACOCKE, *What is a logical constant?*, ***Journal of Philosophy***, vol. 73 (1976), pp. 221–40.

[35] H. PUTNAM, *Mathematics without foundations*, 1967, ***Mathematics, matter and method: Philosophical papers I***, Cambridge, 1975, pp. 43–59.

[36] ———, *Philosophy of logic*, 1971, ***Mathematics, matter and method: Philosophical papers I***, Cambridge, 1975, pp. 323–57.

[37] W. V. QUINE, ***Word and object***, MIT, Cambridge, 1960.

[38] ———, *Existence and quantification*, ***Ontological relativity and other essays***, Columbia, 1969, pp. 91–113.

[39] ———, ***Philosophy of logic***, Prentice Hall, Englewood Cliffs, 1970.

[40] GREG RAY, *Logical consequence: A defense of Tarski*, to appear.

[41] M. D. RESNIK, *Mathematics as a science of patterns: Ontology and reference*, ***Noûs***, vol. 15 (1981), pp. 529–50.

[42] B. RUSSELL, ***Introduction to mathematical philosophy***, Allen and Unwin, London, 1919.

[43] ———, ***The principles of mathematics***, second ed., Norton, New York, 1938.

[44] O. SHAGRIR, *Computation and its relevance to cognition*, ***Ph.D. thesis***, UCSD, 1994.

[45] S. SHAPIRO, ***Foundations without foundationalism: A case for second-order l***, Oxford, 1991.

[46] G. SHER, *A conception of Tarskian logic*, ***Pacific Philosophical Quarterly***, vol. 70 (1989), pp. 341–69.

[47] ———, ***The bounds of logic: A generalized viewpoint***, Bradford, MIT, Cambridge, 1991.

[48] ———, *A new solution to the problem of truth (abstract)*, ***Bulletin of Symbolic Logic***, vol. 1 (1995), p. 131.

[49] ———, *On the possibility of a substantive theory of truth*, manuscript, 1996.

[50] ———, *Semantics and logic*, ***Handbook of contemporary semantic theory*** (S. Lappin, editor), Blackwell, Oxford, 1996, pp. 511–37.

[51] P. SIMONS, *Bolzano, Tarski, and the limits of logic*, ***Philosophia Naturalis***, vol. 24 (1987), pp. 378–405.

[52] A. TARSKI, *The concept of truth in formalized languages*, 1933, in Tarski [1983], pp. 152–278.

[53] ———, *Some observations on the concepts of ω-consistency and ω-completeness*, 1933, in Tarski [1983], pp. 279–95.

[54] ———, *The establishment of scientific semantics*, 1936, in Tarski [1983], pp. 401–8.

[55] ———, *On the concept of logical consequence*, 1936, in Tarski [1983], pp. 409–20.

[56] ———, (J. Corcoran, editor), Hackett, second ed., 1983, translated by Woodger.

[57] ———, *What are logical notions?*, 1966, ***History and philosophy of logic*** *7* (J. Corcoran, editor), 1986, pp. 143–54.

[58] J. VAN BENTHEM, ***Essays in logical semantics***, Reidel, Dordrecht, 1986.

[59] ———, *Polyadic quantifiers*, ***Linguistics and Philosophy***, vol. 12 (1989), pp. 437–64.

[60] R. L. VAUGHT, *Model theory before 1945*, ***Proceedings of the Tarski Symposium*** (Henkin et al., editors), AMS, 1974, pp. 153–72.

[61] D. WESTERSTÅHL, *Quantifiers in formal and natural languages*, ***Handbook of philosophical logic*** (Gabbay and Guenthner, editors), vol. 4, Reidel, Dordrecht, 1989.

[62] L. WITTGENSTEIN, ***Tractatus logico-philosophicus***, Routledge and Paul, 1921, translated by Pears and McGuinness, 1961.

DEPARTMENT OF PHILOSOPHY
THE UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CA 92093-0302, USA

E-mail: gsher@ucsd.edu