Notre Dame Journal of Formal Logic
Volume 41, Number 1, 2000

BOOK REVIEW

Scott Soames. *Understanding Truth*. Oxford University Press, New York, 1998. ix+268 pages.

1. Introduction

Scott Soames's book, *Understanding Truth*, offers a skillfully written introduction to a wide range of problems of contemporary interest involving the notion of *truth*. This range includes the question of the bearers of truth values, Tarski's definition of truth, the Liar Paradox and solutions to it, vagueness and the sorites paradox, and deflationism. Throughout, the book is well written and closely argued, and Soames's judgments are judicious and generally sound. This would be an ideal introduction to the subject for advanced undergraduate and graduate students were it not for a single flaw: the book fails to refer to much important work on the subject done in the late '80s and early '90s. I will make note of some of the more important lacunae below, and I hope that future additions of the book will fill these bibliographical holes. Otherwise, I am worried that philosophers relying on Soames's text for a survey of the state of the art in 1999 will overlook some very significant books and articles, works which are in many cases directly relevant to the issues Soames discusses.

The book comprises three parts: Foundational Issues, Two Theories of Truth, and Extensions. In the first part, Soames addresses two questions: What are the bearers of truth? and Does truth have a significant role to play in philosophical theories? In the second part, Soames deals with the formal theories of truth developed by Tarski and Kripke, and he develops a positive case for his own interpretation of Kripke's theory in which 'truth' is treated as a partially defined predicate. Finally, in the third part, Soames turns to the problems of vagueness and the sorites paradox and to the issues raised by deflationary theories of truth.

2. Truth Bearers

In Chapter 1 Soames offers a series of convincing arguments in favor of the thesis that the fundamental bearers of the property of truth are propositions. An eternal sentence (one free of indexicals, demonstratives, and significant uses of tense) is true just in

Received August 25, 2001; printed July 15, 2002
2001 Mathematics Subject Classification: Primary, 03B65; Secondary, 03B52
Keywords: truth, Tarski, liar paradox, vagueness, deflationism

case it expresses a true proposition. In the more general case, a sentence meaning can be thought of as a function from contexts to propositions. Hence, a sentence uttered in a context is true just in case the sentence's meaning maps that context onto a true proposition.

Soames doesn't offer a substantive theory of propositions. Are propositions merely sets of worlds, as Stalnaker and Lewis would have it, or are they structured objects (some sort of complex of objects, properties, and relations), as they are according to Russell's account? Nor does he address the issue raised by Barwise and Etchemendy in *The Liar* [1]: Do propositions consist merely of a complex of objects, properties, and relations, or do they always also contain some reference situation, some actual part of the world which they must be taken to be about? These are issues quite central to the nature of truth and the lack of attention to them leaves a significant gap in Soames's picture.

In addition, Soames does not consider the possibility that limiting our attention to the truth of propositions may deprive us of important insights. According to many contemporary accounts of truth and representation, a *true* statement is one that fulfills the proper function of a statement, that is, that faithfully meets our legitimate expectations of statements. If this is right, then a theory of truth (as applied to statements, eternal sentences, and beliefs) must be a theory of representation: an account of the proper functions of these representations. By focusing instead on the philosophical artifact of *true propositions*, Soames assumes that the theory of truth is separable from a theory of representation: the theory of representation deals with how statements and sentences express propositions whereas the theory of truth investigates the difference between true and false propositions. As Soames's book in the end demonstrates, this division of labor is infelicitous. Once the theory of representation has been hived off, there is nothing interesting left for the theory of truth to do.

3. Skepticism about Truth

Soames discusses several versions of skepticism about theories of truth in Chapter 2: Frege's argument that truth is indefinable, the logical positivists' claim that truth is irrelevant because epistemically inaccessible, the objection to classical (correspondence) theories of truth that they are excessively metaphysical, the claim by redundancy theorists that truth is trivial, and Tarski's argument that our ordinary conception of truth is inconsistent (as shown by the Liar Paradox).

Frege argued that any definition is comprehensible only to someone with a prior grasp of the concept of truth, since a definition is (qua speech act) a stipulation that a certain biconditional shall always be taken as true. The cogency of Frege's argument depends crucially on what one takes to be the criteria of a successful definition. If it is sufficient that the definiendum be coextensive with the definiens, or even that this coextensiveness hold as a matter of metaphysical necessity, then Frege's argument fails to rule out the possibility of a successful definition of truth. However, if a successful definition must be capable, at least in principle, of introducing the concept of the definiendum to a thinker with no prior grasp of that concept, then Frege's argument is conclusive.

As Soames puts it: "Thus if truth is definable, the notion of truth so defined must be one that everyone already grasps in the absence of any definition" (p. 29). Soames assumes that there is nothing problematic about a definition whose comprehensibility presupposes a prior grasp of the definiendum. It would seem inappropriate to label such a biconditional a *definition* of truth: something like *explication* or *elucidation* would seem to be more appropriate.

Soames next turns to the argument made by many logical positivists (Soames takes Reichenbach as a representative example) that truth is philosophically useless because, scientifically speaking, truth can never be established with absolute certainty. Reichenbach consequently labels truth a "fictive" notion ([18], pp. 187–8). As Soames points out, this seems to involve nothing more than a confusion of truth and certainty. If Reichenbach's argument were successful, it would invalidate the use of every concept, including *gram* or *meter*, since we cannot establish any measurement with absolute certainty.

The third objection to truth is that of the antirealists of Dummett's school. Here the objection is to a supposedly classical or correspondence theory of truth as used to underwrite classical logic, especially the law of excluded middle. As Soames sees it, this dispute between realists and antirealists is a dispute about *meaning* and not at all about *truth*. Dummett objects to a truth-conditional semantics only insofar as it includes two critical features:

1. it conceives of the meanings of negation and disjunction in a way that validates the law of excluded middle; and
2. it is willing to treat unverifiable sentences as expressing genuine propositions.

Thus, neither of the crucial issues concerns truth per se: one concerns the interpretation of the sentential connectives and the second concerns the connections between sentences and propositions. In addition, one could hold a realist (bivalent) conception of truth while rejecting truth-conditional theories of meaning (as Soames himself does). Thus, Soames argues, the realism/antirealism dispute has nothing to do with the theory of truth.

These points have some merit, but they reinforce my worry that Soames has defined the theory of truth so narrowly as to exclude any problem of philosophical significance. The question of whether there exist propositions capable of unverifiable truth seems to have *something* to do with the nature of truth, as does the question of whether the truth-values of the logically complex depend on the truth-values of their parts in accordance with the standard, classical account.

In response to the attempt by redundancy theorists to trivialize the notion of truth, Soames makes a number of compelling points. However, he fails to cite Gupta's important paper "A critique of deflationism" [8] in which similar objections were lodged.

Redundancy theorists claim that a statement S and the statement 'S is true' express the very same proposition. Redundancy accounts face an immediate problem: how to represent the content of such uses of 'true' as this: 'Everything Mary asserted is true'. There is no statement explicitly mentioned in this sentence as being true, so it is not immediately clear how to eliminate the occurrence of the supposedly redundant 'true'. Redundancy theorists have traditionally resorted to second-order quantification:

$$(P)(\text{if Mary asserted that } P, \text{then } P).$$

The second-order quantification can be interpreted either substitutionally or quantificationally. Soames rehearses most of the standard objections to substitutional quantification (a better summary of which is provided by Gupta). Unfortunately, he fails to discuss Grover's influential *prosentential* version of substitutional quantification and thus leaves this account unchallenged. This prosentential account has recently been defended in some detail by Brandom in [2].

As Soames points out, an objectual interpretation of second-order quantification is unattractive to redundancy theorists since it would commit them to propositions as members of their ontology. If propositions exist, and some of them are true and some not, then it is hard to resist the conclusion that 'is true' represents some kind of property of propositions, which would mean that 'S is true' does, while S itself does not, involve this special property.

Another challenge to truth comes from the Liar Paradox. Tarski took the semantical paradoxes to show that natural language is *inconsistent* because of its inclusion of an unrestricted truth predicate. What does it mean to say that natural language is 'inconsistent'? As Soames informs us, Salmon has suggested that an inconsistent language contains at least one contradiction that is true. Under plausible assumptions about English, we can derive that both the Liar L and its negation are true.

$$(L)\ L \text{ is not true.}$$

Soames argues that the conclusion that both L and its negation are true is clearly unacceptable since it is tantamount to accepting a contradiction ourselves. Consequently, Soames urges that we give up instead one of the Tarskian assumptions leading to this conclusion. This problem Soames takes up in Chapters 5 and 6.

4. Tarski's Theory of Truth

In Chapter 3 Soames offers an exemplary exposition of Tarski's definition of truth. He well motivates Tarski's use of infinite sequences of objects in his definition of truth for the quantifiers and he illuminates the connection between Tarski's theory and Gödel's first incompleteness theorem.

Soames also provides an interesting appendix on the ever-thorny and often-surprising problem of developing a consistent theory of quotation and quotation-terms. Tarski objected to any sort of quantification into direct-quotational contexts as in the following abortive definition of truth:

For all S, 'S' is true in L iff S.

Tarski insisted that quotational terms are syntactically simple so that the above clause should be interpreted as

For all S, the thirteenth letter of the alphabet is true iff S.

Soames argues that, although Tarski's account of quotation is acceptable, there is a viable alternative. In Soames's alternative, 'Quote' is a one-place functor. For any α, where α is a term, predicate, sentence, or other syntactic structure, 'Quote(α)' is a term that denotes α. Now, if S is interpreted as a substitutional variable, the definition rejected by Tarski is now well formed:

For all S, 'S' is true in L iff S.

The Liar Paradox will precipitate a contradiction in Soames's theory unless the substitutional variable S is stipulated to belong only to the metalanguage, not to the object language L for which truth is being defined.

Could such an account provide a noncircular theory of truth? Platts [16] argued that substitutional quantification presupposes the notion of truth since the interpretation of

$\exists S \Phi(S)$, where S is a substitutional variable, is as follows: there is a name (sentence, open formula, or what have you) that, when replacing S in $\Phi(S)$, results in a *true* sentence. Soames offers (very tentatively) one or two quite weak objections each of which seems to beg the question. In the end, he concludes with the tautologous remark that Platts is wrong if substitutional quantification can be taken as a semantic primitive. Soames is "inclined to think that it is," but he offers nothing that would dissuade me from the opposite inclination.

Soames takes up the philosophical significance of Tarski's theory in Chapter 4. He begins by asking whether Tarski's theory is an analysis of the English predicate 'is true'. He argues that it is clearly not for two reasons: (1) it applies only to sentences, not to propositions, and (2) it is limited to a single language, English, whereas 'is true' in English can be applied to other languages. Of course, Tarski doesn't claim to provide such an analysis since he was convinced that the Liar Paradox demonstrated that our ordinary notion of truth is radically defective. Tarski's definition, or *explication*, of truth is intended, not as an account of how 'is true' is to be understood, but of how it must be understood, if it is to fulfill its most central functions while avoiding contradiction. Thus, Tarski's definition is, as Soames recognizes, explicitly revisionary.

As Soames points out, the theoretical fruitfulness of Tarski's theory is beyond doubt. All of contemporary model theory is merely a footnote to Tarski's work.

Soames discusses two possible tensions between Tarski's definition of truth and Davidson's theory of our knowledge of linguistic meaning both in footnote 32 to Chapter 3 and in the body of Chapter 4. Davidson proposed that to understand the meaning of a sentence it suffices to know that sentence's truth-conditions as they would be generated by a Tarski-like truth theory for the language to which the sentence belongs. In the footnote to Chapter 3, Soames argues that this Davidsonian thesis is incompatible with Tarski's claim that his theory provides a *definition* of truth.

According to Soames, definitions are knowable a priori. Hence, if Tarski's theory provides a definition of truth, each of its clauses must be knowable a priori. However, according to Davidson, a theory of meaning must be empirical not a priori. The axioms of a truth theory for a language are empirically confirmed by observing the linguistic behavior of the relevant community.

This argument for incompatibility seems sound, once one grants Soames's assumptions that successful definitions are a priori knowable. However, a priori equivalence is a very high standard for definition to reach, much higher than Tarski claims for his own definition. In their book on circularity [9], Gupta and Belnap have argued, quite plausibly, that a much weaker condition, that of intensional equivalence, is sufficient. If one were to adopt the Gupta/Belnap account of definition, it is not at all clear that one couldn't combine a Tarski definition of truth with a Davidsonian account of the knowledge of meanings.

Soames returns to the question of the relationship between Tarski's theory and Davidson's in Chapter 4. He points out that the two followed mutually converse methods: Tarski took meaning for granted to explicate truth and Davidson took truth for granted to explicate meaning. That combining these two approaches would involve a vicious circularity is something Soames takes for granted, but, in light of the work of Gupta and Belnap on the viability of circular definitions, this is a matter that deserves further investigation.

Soames endorses the criticism of Davidson's account of meaning epistemology that has been put forward by Dummett, Evans, McDowell, and others, namely, that Davidson erred in assuming that he could make use of Tarski's technical definition of true-in-L in giving an account of what is required to understand the language L. Soames presses home quite convincingly the claim that a Davidsonian theory of the knowledge of meaning would actually have to make use of our *ordinary* conception of truth, one that applies across languages and to propositions (objects of belief and other attitudes) as well. As Soames points out, Davidson himself came to accept this point.

Soames then turns to Field's critique of Tarski's definition from the point of view of a strict physicalism. Field argues that making the notion of truth physicalistically acceptable was at least one of Tarski's purposes, but it is clear that Tarski's conception of physicalism was more expansive than Field's. Tarski intended to show that semantical propositions could be defined by reference to both narrowly physical and logico-mathematical facts.

Field accepts that Tarski's definition was extensionally correct, but he argues that this is insufficient to ground the claim that semantical facts have been reduced to or defined in terms of physical facts alone. In particular, it is the base clauses of Tarski's definition, for example, that 'snow' is true of x if and only if x is snow, that fall short of what's required, offering only *pseudoreductions* of the semantical terms. What Field thinks is needed is a causal-historical account that explains how the extensions of terms in a language are fixed by the behaviors and environments of language-users.

Soames points out that Field has seriously underestimated the scope of his critique. It is not only the base clauses, but also the recursive clauses of Tarski's definitions (fixing the meanings of the logical constants) that fall short of Field's standards of physicalistic reduction. To say that 'φ or ψ' is true if and only if either φ is true or ψ is true also fails, just as the base clauses do, to explain how the meanings of these constants are fixed by actual linguistic practices. Field cannot possibly offer a causal-historical account of the reference of the logical constants, since according to his sort of physicalism, logical facts are causally inert. A Platonist like myself might well think that a reduction of semantic facts to nonsemantic facts (including causally active logico-mathematical facts) along the lines suggested by Field would be a worthwhile endeavor. However, such a position is unavailable to Field.

As Soames argues, another difficulty for Field lies in the fact that, whereas for Tarski, the bearers of truth were sentences (that is, sentence-types), for Field they are sentence-tokens. It would seem that the success of Field's reduction would depend on the very dubious assumption that enough concrete tokens actually exist. For example, suppose that token A is a token of the sentence φ but the sentence-type φ has never actually been tokened at all. (In a language such as English, it might be plausible to assume that φ would be tokened by some part of A but there could certainly be a language in which negation is not expressed by adding some particle or particles to the unnegated base.) In such a case, Field's version of the recursive clause will fail to be extensionally correct: A might be true even though there is no token B to which A bears the negation-relation and which is itself not true.

In the Appendix to Chapter 4, Soames takes up the critique of Tarski's definition of logical consequence lodged by Etchemendy in his book *The Concept of Logical Consequence* [4]. In his response to Etchemendy's critique, Soames relies heavily on

Gomez-Torrente's 1996 Princeton dissertation, "Tarski's Definition of Logical Consequence: Historical and Philosophical Aspects" [7]. However, Soames seems to be entirely unaware of Ray's penetrating and very influential response to Etchemendy, "Logical consequence: A defense of Tarski" [17] which appeared also in 1996. I hope that this oversight will not lead to an unfortunate neglect of Ray's article. Ray's response to Etchemendy at some points parallels that of Soames (especially on the question of interpreting counterfactuals involving the supposition of a finite mathematical universe) and at many points offers complementary arguments.

In a 1935 lecture, first published in 1936 under the title "On the concept of logical consequence" [22], Tarski proposed to identify the logical truth of a sentence with the truth of that sentence under all possible substitutions performed on the nonlogical elements of the sentence. This is similar to, but as Etchemendy points out, subtly different from the modern model-theoretic characterization of logical truth as truth in all models, that is, truth under all possible *interpretations* of the nonlogical elements of the sentence. As Etchemendy asserts, Tarski's 1935 account seems to take the domain of actually existing things as fixed, varying the extension of names and predicates within that domain, whereas the modern model-theoretic characterization insists upon varying the domain of quantification from one model to the next. What is controversial in Etchemendy's interpretation is his claim that this peculiar feature of Tarski's paper constituted a deep and serious part of Tarski's project. Both Ray and Gomez-Torrente dispute Etchemendy's claim, arguing that Tarski must in fact have had something like the modern account in mind and that any appearance to the contrary is simply the product of Tarski's attempt to simplify his account for the benefit of a logically unsophisticated audience.

Etchemendy's critique of Tarski is largely independent of this hermeneutic dispute since it will apply with nearly equal force to a proposed analysis of logical truth as truth in all models. Etchemendy has two sets of arguments against the adequacy of this model-theoretic analysis: one set concerns first-order logic and the other second-order logic. In the case of first-order logic, Etchemendy points out that there are first-order sentences that are not logically true but that are true in all *finite* models. Etchemendy points out that the material adequacy of the model-theoretic definition of logical truth depends upon the actual existence of models with infinite domains. Were the mathematical universe sufficiently impoverished (as mathematical finitists believe it to be in fact), there would not exist mathematical structures large enough to serve as infinite models and the proposed definition would fail. So far, Etchemendy's claims are uncontroversial.

Etchemendy proceeds to argue that the model-theoretic account can work as an *analysis* of logical truth only if its ontological presuppositions (that is, the existence of infinitely large structures) are themselves logically or analytically true. Yet, according to the standard view of these things, claims about the existence of infinitely large structures are not themselves logically true. If they were, then the very sentences mentioned above (sentences true in all but only finite models) would be wrongly characterized as logically false.

Clearly, the most contentious claim in Etchemendy's account is his principle, which I will call 'Etchemendy's principle', to the effect that the ontological presuppositions of a correct analysis must themselves be analytically true. This principle is disputed by Ray, Gomez-Torrente, and Soames. Historically speaking, it seems very unlikely that Tarski himself would have embraced Etchemendy's principle. He accepts no more stringent requirement for his definitions, including his definition of truth, than that of *material adequacy*, and Etchemendy's argument concerning first-order logic does not dispute the material adequacy of his definition of logical truth, given the actual infinity of the mathematical universe.

Even if we set the bar higher and require a correct analysis to meet the standard of *necessary coextensiveness* of the definiendum and the definiens, it would seem that the model-theoretic analysis is under no serious threat since Etchemendy offers no argument for thinking that the mathematical universe is *contingently* finite, only that its infinity is a synthetic, rather than analytic, truth (a point made by Ray, as well as Soames). As Soames points out, Etchemendy offers no argument for his principle. Soames attempts to reconstruct an argument along these lines:

S1 If the MT analysis is correct, then for any sentence *S*, the claim that *S* is a logical truth is analytically equivalent to the claim that *S* is true in all models.

S2 It is analytic that a certain sentence *F* is true in all finite models (where *F* is a sentence true in all finite models but false in some infinite ones).

LC What follows logically from analytic truths is itself analytic.

S3 So, if the MT analysis is correct, it must be an analytic truth that if *F* is not logically true, then there must be infinite models (from S1, S2, and LC).

S4 *F* is not logically true (by common agreement as well as according to the MT account).

S5b If *S* is not a logical truth, then it is analytic that *S* is not a logical truth.

7 So, it is analytic that *F* is not a logical truth (from S4, S5b).

S6 So, if the MT analysis is correct, it is analytic that there are infinite models (from S3, 7, and LC).

Although the definition of analytic truths is notoriously difficult, this argument seems to be sound with the possible exception of principle S5b. In addition, Soames argues that the conclusion is not a genuine reductio since it does not assert that it is a *logical truth* that there are infinite models but only that it is *analytically true* that there are.

Soames points out that it is not obvious what defense Etchemendy could give for S5b. Since first-order logic is complete but not decidable, there is an argument, for the correctness of a complementary principle, S5a, that is not available in the case of S5b.

S5a If *S* is a logical truth, then it is analytic that *S* is a logical truth.

If some sound and complete logical calculus is such that it is analytic that all of its axioms are logically true, then principle LC would entail S5a. However, since first-order logic is incomplete, there is no effective procedure that would establish the logical contingency of all logically contingent sentences and so, obviously there is no effective procedure such that it is analytically true that it will do so.

Nonetheless, if the truth of S5a is conceded, there is a strong case to be made for S5b. Consider what would have to be the case for S5a to be true but S5b to be false. It would mean that there was a predicate, 'is logically true', that would be *analytically* true of sentences, whenever it is true of them, but *synthetically* false of the remainder. If statements asserting logical truth are analytically true, then they are (in some sense) *made true* by facts about the meanings of their elements and their logical form. Sentences affirming logical truth can, if S5a is true, be made true by nothing else. Hence, if they are *not* made true by their meanings and logical form, they are false, and false, it would seem, by virtue of the fact that their meanings and logical form do not make them analytically true. Hence, it would seem to follow that they are made false by virtue of the meanings of their elements and their logical form, and so are analytically false, vindicating S5b.

I think there may well be something deeply defective about the categories *analytic/synthetic* and *logically true/substantively true*. However, Soames himself accepts these categories as essentially unproblematic. If these categories are unproblematic, it would seem that Etchemendy would have a strong case for S5b and thus a much stronger argument than Soames recognizes.

In fact, as Soames recognizes, it is reasonably clear that Tarski would object to principle S1, that is, Tarski would not claim that his proposed definition is analytically true (see Ray's paper for substantiation of this point). For Tarski, it was enough for his definitions to be *materially adequate*. This is far too low a standard to accept, however. A good definition ought to secure at least necessary coextensiveness. Gupta and Belnap make a good case for an even stronger condition which they label 'intensional equivalence'. However, even intensional equivalence is considerably weaker than *analytic equivalence* which is what Etchemendy's first-order argument requires.

For this reason, Etchemendy's argument concerning second-order logic is more pertinent to Tarski's project since in it Etchemendy argues that Tarski's definition of logical truth is not even materially adequate. Etchemendy's argument depends on the fact that we can construct second-order sentences *A* and *B* of such a kind that *A* is true in all possible models (all models constructible from sets) if and only if the continuum hypothesis is true, while *B* is true in all models if and only if the continuum hypothesis is false. As Gomez-Torrente reconstructs the argument, it runs as follows:

S1 *A* is true in all models iff CH is true.

S2 *B* is true in all models iff CH is false.

S3 Neither CH nor its negation is logically true.

S4 So, neither the claim that *A* is true in all models nor the claim that *B* is true in all models is logically true.

S5 For all *S*, if *S* is a logical truth, then the claim that *S* is true in all models is a logical truth.

S6 Thus, neither *A* nor *B* is a logical truth.

S7 Either CH or not CH.

S8 Either *A* is true in all models or *B* is true in all models.

S9 If the model-theoretic analysis of logical truth is correct, then either *A* or *B* is a logical truth.

S10 So, the model-theoretic analysis is incorrect.

As Soames points out, this reconstruction has an obvious flaw: S4 does not follow from S1 through S3. Instead, Soames argues, it follows from S3, together with the claims that S1 and S2 are logical truths. This is still not quite right: in addition to S3 and the logical truth of S1 and S2, we need an analogue of LC above:

LE Whatever is logically equivalent to a logical truth is itself a logical truth.

Soames then offers two further reconstructions of this argument neither of which is ultimately successful. However, this critique of Etchemendy's argument suffers from a critical flaw—nothing resembling the subargument S1 – S6 can be found in Etchemendy's book. Instead, Etchemendy offers two quite different arguments concerning second-order logic and the continuum hypothesis. First (on page 123), he observes that if the continuum hypothesis is true in fact, then it will be true in every (standard) model of set theory since it will be impossible to falsify the hypothesis by building a model with a cardinal number between that of the natural numbers and that of the real numbers, there being no such set available for the construction of the model. Hence, if the continuum hypothesis is true, and the model-theoretic analysis of logical truth is correct, it follows that the continuum hypothesis is not only true but *logically* so. Etchemendy takes it as obvious that the continuum hypothesis is not logically true (an intuition that I and many others share) from which it follows that if the continuum hypothesis is true, Tarski's definition is not materially adequate.

A defender of the model-theoretic analysis could insist that this observation gives us compelling reason to hold the continuum hypothesis to be false (and, more generally, to embrace a kind of principle of plenitude in set theory). However, this suggestion seems unsound: surely we should continue to consider the continuum hypothesis as unsettled and so we should consider the question of the material adequacy of the model-theoretic analysis for second-order logic to be likewise an open question.

Etchemendy has a second argument which is somewhat closer to Gomez-Torrente's reconstruction. Etchemendy claims that he can construct two sentences *A* and *B* that will be true in all models if the continuum hypothesis is true or false, respectively. In the first case this is straightforward and unobjectionable: the continuum hypothesis can simply be stated in the language of second-order logic, and if the hypothesis is true, this second-order statement will be true in any standard model, that is, a model in which ϵ receives its intended interpretation. In the case of *B*, we can construct a sentence that says that *if* there exists a set with the structure of the real numbers, then there exists a set whose cardinality is intermediate between that of the natural numbers and the hypothesized set of real numbers. If the continuum hypothesis is false, that is, if there exist such intermediate cardinals, then any standard, full model will validate *B*, vindicating Etchemendy's claim.

Unlike the reconstruction offered by Gomez-Torrente, Etchemendy offers no problematic proof that *A* and *B* are not logically true. He takes it as obvious that they are not. Sentence *A* is simply the statement of the continuum hypothesis in a second-order language: if the continuum hypothesis is not logically true, neither can *A* be. Sentence *B* is the claim that if there is a set with the structure of the real numbers, then there is a set whose cardinality is intermediate between that of the first set and the cardinality of some set with the structure of the natural numbers. This is not exactly the negation of the continuum hypothesis but it is clearly a substantive claim about the set-theoretic universe. Sentence *B* is essentially the disjunction 'either there is no set with the structure of the real numbers or not-CH'. It's hard to see how *B* could be logically true, and Soames offers no explanation of its logical truth to dispel this impression. Hence, Etchemendy's argument is considerably simpler and less problematic than Gomez-Torrente's reconstruction:

E1 ($\approx$ S1) *A* is true in every standard model iff CH is true.
E2 ($\approx$ S2) *B* is true in every standard, full model iff CH is false.
E3 If the MT analysis is correct, and if that analysis is committed to the use of standard, full models, then either *A* or *B* is logically true.
E4 ($=$ S6) However, neither *A* nor *B* is logically true.
E5 Therefore, if the MT analysis is committed to the use of standard, full models, then it is not even materially adequate.

Etchemendy recognizes that the defender of the model-theoretic analysis can always save the phenomena by introducing more models. For example, if we were to allow the use of so-called general or Henkin models (models in which the domain of second-order quantification is a proper subset of the powerset of the first-order domain), then we can build models that falsify *B* even if the continuum hypothesis is false. If we allow ourselves nonstandard models, models in which the interpretation of ϵ does not correspond to the actual membership relation among sets in the second-order domain, we can falsify *A* even if the continuum hypothesis is true. However, the move even to general models comes at a heavy cost: second-order logical truth essentially collapses into first-order logical truth (the exceptions are relatively trivial). For those who are skeptical about the logical status of second-order logic (like Quine), this may be a welcome result, but for many of us it is strongly counterintuitive.

This illustrates a wider theme in Etchemendy's book that Soames neglects. As Etchemendy repeatedly points out, it is always possible for the Tarskian to achieve material adequacy for the definition by simply gerrymandering the set of *acceptable* models. If the definition threatens to overgenerate logical truths, the Tarskian must simply add new models of the appropriate kind; if it threatens to undergenerate, the Tarskian must simply exclude some models. Etchemendy is not claiming that it is impossible for the Tarskian to achieve material adequacy by such moves, but instead he is claiming that the unprincipled nature of these alterations reveals that the Tarskian analysis is not informative or substantive. Model theory merely serves to represent or systematize our intuitions about logical truth: it reveals nothing about the *essence* of logical truth or logical consequence.

In fact, the bulk of Etchemendy's book is devoted to the issue of *cross-term restrictions*, which Etchemendy takes to be a paradigm example of the unprincipled nature of the Tarskian definition. Unlike Ray, Soames nowhere engages this part of Etchemendy's argument, leaving his critique of Etchemendy's arguments seriously incomplete.

5. The Liar Paradox

In Chapter 6 Soames deals with some recent work on the Liar paradox, especially Burge's 1979 paper, "Semantical paradox" [3]. Once again, Soames overlooks a great deal of work, inspired by Burge's paper, that has been done over the last fifteen years. This work includes books by Barwise and Etchemendy [1], Simmons [20], and the reviewer (Koons [10]) as well as an article by Gaifman [5]. This is significant because many of the objections to Burge's contextual/hierarchical approach that Soames raises have already been extensively discussed in the literature.

Burge's approach involves the claim that the interpretation of 'true' varies from context to context. This enables Burge to argue that the Liar sentence is in fact *true*, albeit true in a different context from the one in which the utterance of 'true' in the Liar occurs. This contextual resolution can be combined with Tarski's idea of a hierarchy of metalanguages resulting in what is known as the *contextual/hierarchical* solution, followed by Burge [3], Koons [10], and Glanzberg [6]. A nonhierarchical version of the contextual resolution has also been offered in recent years by Gaifman [5], Barwise and Etchemendy [1], Simmons [20], and Koons [11].

On the hierarchical/contextual account, each use of the predicate 'true' is to be assigned an ordinal number, corresponding to a position in a hierarchy of interpretations. So, for instance, a ground-level Liar sentence could be interpreted as saying 'This sentence is not true_0'. The interpretation of 'true_0' can be generated in a variety of ways: a popular way is to use the minimum fixed point of Kripke in [12]. At this fixed point, the Liar is neither in the extension nor the antiextension of 'true_0': it suffers a truth_0-value gap. However, there are additional interpretations of 'true'. For example, to arrive at the interpretation of 'true_1', we *close off* (to use Kripke's phrase) the interpretation of 'true_0', putting the Liar sentence into the antiextension of 'true_0', since it fails to be true_0. Since the Liar sentence says that it is *not* $\textit{true}_0$, the Liar sentence itself goes into the extension of 'true_1'. This explains why it is *correct* to observe that the Liar sentence is not true. This move enables the contextual account to avoid the so-called Strengthened Liar paradox.

On a nonhierarchical but contextual account, the interpretation of 'true' shifts from context to context but the varying interpretations do not form an ordering of any kind. For example, in [20] and [11], following a suggestion by Gödel, most everyday interpretations of 'true' receive a bivalent interpretation, with every meaningful utterance going either into the extension or antiextension. It is only exceptional utterances, like utterances of the Liar, the truth-teller, and other problematically circular or ungrounded contents, that receive partial, gappy interpretations.

Soames seems to be unaware of the existence of nonhierarchical versions of the contextual solution to the Liar, but he does offer four objections to the contextual/hierarchical account. First, he claims that many unparadoxical things are inexpressible in a hierarchy of languages. Moreover, as Soames points out, many of the English sentences needed to define, set up, or describe the hierarchy have this level-transcending character. I discussed this problem in my 1992 book, arguing that hierarchy-transcending statements should be taken as *schematic* in nature, following up on a suggestion by Burge. In addition, this objection does not apply to nonhierarchical versions of the contextual approach.

Second, Soames argues that the contextual approach cannot be generalized to cover uses of 'true' to evaluate sentences in other languages. Soames envisages the hierarchical approach as committed to separate hierarchies for each natural language, English, Dutch, and so on. It is hard to see why the hierarchical approach need take this form. Why couldn't all natural languages share a common hierarchy? Or, if one prefers, one could talk about a hierarchy of propositions, with each level introducing a new interpretation of the *truth*-concept and with propositions as the primary truth-bearers. Nothing about the contextual account forces a narrowly linguistic view of the nature of truth. Barwise and Etchemendy's account, for example, is couched entirely in terms of propositions.

Third, Soames argues that the contextual account yields an unacceptable fragmenting of the meaning of 'true'. On the contextual approach, he argues, the word 'true' is infinitely ambiguous, in effect, comprising an infinite collection of homonyms, raising questions about how such a word (or infinite collection of words) is learnable. Soames seems to ignore here the analogy that contextual theorists, following Burge, have offered from the very beginning. The word 'here' varies in interpretation from one context to the next but does not become ambiguous or unlearnable thereby. To use Kaplan's terminology, the word 'true' possesses a constant *character*, even though its *content* varies from one use to the next.

Fourth, Soames argues that it is an impossible task to assign the appropriate index to each use of 'true'. Soames refers to cases introduced by Kripke in [12] in which the paradoxicality of certain statements, such as 'Everything Dean said about Watergate is false', turns on empirical facts that may be unknown to the speaker and his audience. This is a problem that has been much discussed since Burge's paper: Gaifman, Simmons, and I all offer detailed solutions none of which is mentioned by Soames.

Soames concludes this chapter by arguing that the contextual solution is unnecessary since the Liar paradox can be avoided simply by giving up the principle of bivalence. Soames here completely ignores the problem of the Strengthened Liar which is a stock-in-trade for contextualist critiques of noncontextual, truth-value gap accounts. Soames argues that the Liar is neither true nor false. This implies that the Liar is not true, which is simply a restatement of the Liar itself. Soames must explain how something that is a logical consequence of his own account could fail (by that very account) to be true.

6. Truth as a Partially Defined Predicate

In Chapter 7 Soames offers an account of 'true' as a partially interpreted predicate, an account which he first proposed in 1989. Here again there is a bibliographical problem. An account very similar to Soames's was developed independently by McGee in [14], a monograph that received the Johnsonian Prize in 1988. In this monograph, McGee goes considerably further than Soames does in filling out the semantical details and demonstrating the relevant metalogical results. Nonetheless, McGee's work is never mentioned.

Soames makes a distinction between the extension of a predicate and the *determinate* extension of the predicate, a distinction identical to the one that plays a central role in McGee's book. When these indeterminacies produce indeterminate truth values, Soames proposes to use strong Kleene three-valued logic to interpret logically complex sentences. Surprisingly, he does not even mention any alternatives, such as the method of supervaluation, which has played such a prominent role in Gupta's work on the Liar. Soames also distinguishes between *rejecting* a proposition and accepting its negation, a distinction first proposed, I think, by Parsons in [15].

Soames proposes using the Kripke construction, in combination with Strong Kleene truth tables, to reach the minimal fixed point. He doesn't explicitly consider the range of alternatives to the fixed point that have been proposed, including maximal fixed points or Barwise's *self-sufficient sets*. In fact, Soames eventually replaces the minimal fixed point with Kripke's *intrinsic truth values*, an interpretation at which some ungrounded sentences receive truth values but only when this can be done without arbitrariness. In other words, ungrounded sentences that could consistently be given one and only one classical truth value are assigned this value.

Soames then takes up the problem of the *resiliency* of the Liar. This is not exactly the same thing as the Strengthened Liar paradox which Soames consistently overlooks. He proposes that the Liar has a certain kind of negative semantic assessment: it lacks a determinate truth value and so should be rejected, not accepted. As Simmons [19] has pointed out, the diagonal argument that generates the Liar can be applied to any semantic assessment whatsoever. Consider, for example, a Strengthened Liar L' consisting of the sentence 'Sentence L' has the must-be-rejected semantic status'. Soames's account entails that L' must be rejected which is exactly the content of L'. Hence, Soames's account logically entails something that, by its own light, must be rejected. This is problematic to say the least.

What Soames calls the *resiliency of the Liar* is a subtly different phenomenon. Soames admits that once we have introduced the status *determinately true*, we create a new conceptual framework within which a new Liar can be constructed, namely, L'', 'the sentence L'' is not determinately true'. Soames does not see this as an objection to his account since he denies that the new Liar L'' was available in natural language in its pristine form, and it is only the function of 'true' in that original context that his account is supposed to illuminate.

First of all, it is not at all clear that the introduction of the notion of *determinate truth* necessitated a kind of conceptual revolution. In fact, the resources for constructing such a notion were fully present in everyday English. Otherwise, it is hard to see how Soames could have succeeded in introducing the notion with so little difficulty. Second, the status of *being rejectable* must have been available in pristine English since this notion is needed to explain the upshot of his account, that is, to explain what exactly his account does with the Liar. As I pointed out, a strengthened Liar can be couched entirely in terms of rejectability without reference to the technical phrase 'determinately true'.

Since Soames does not in the end avoid a Tarski-like hierarchy (in his case, it takes the form of the sequence 'true', 'determinately true', 'determinately determinately true', . . .) and since Soames's principal objection to the contextual/hierarchical solution was to the restrictions that such a hierarchy introduces, why should his account be preferred? Soames offers three reasons.

1. His theory provides "a plausible account of how the notion of truth might be introduced into a language and understood by its speakers." (p. 181)
2. "It makes room for a lot of expressive power at the very first level of the hierarchy."
3. His theory provides a model for explaining how the languages in the hierarchy are successively generated.

I cannot see how any of these reasons give an advantage to his account over that of the contextual school. Contextualists, following Burge, have made use of the same sort of Kripkean inductive construction that generates virtues (1) and (3) on Soames's list, and nothing prevents a contextualist from using the same (and even better) techniques to enrich the expressive power of each level, including the first one (again, see [20] and [11] for details). On the other side of the ledger, the contextualists avoid the Strengthened Liar phenomenon while Soames does not. In addition, the contextualist solution generalizes to near neighbors of the Liar (including the Knower paradox and related paradoxes involving provability and rational belief), while truth-value gap theories like Soames's do not, as I argued in [10].

7. Vagueness and the Sorites Paradox

In Chapter 7 Soames applies the account of semantic partiality he developed in Chapter 6 to the problems of vagueness and the Sorites paradox. Here again there are serious bibliographical lacunae, especially Soames's neglect of McGee [14] which carries out exactly the same program with a great deal more technical sophistication. In addition, Soames ignores entirely the recent revival of epistemicist accounts of vagueness championed by Sorensen [21] and Williamson [23].

Epistemicists such as Sorensen and Williamson will argue it is hard to see why one would prefer a semantic partiality account of the kind Soames proposes to the epistemicist account, according to which there are sharp boundary lines governing vague predicates of which we are ignorant. Soames explicitly rejects the following principle:

> (V) For any two perceptually indistinguishable patches of color x and y and acceptable context of utterance C, the standards governing the predicate *looks green* cannot include x in the (determinate-) extension of the predicate and y in its (determinate-) anti-extension. (p. 223, note 11)

If V is rejected, then what principled grounds can Soames have for the rejection of the epistemicist account? The epistemicist account has the advantage of preserving classical logic and semantic bivalence. Its only disadvantage is its rejection of principles such as V which commits the epistemicist to the surprising claim that our linguistic practices can establish semantic rules that outrun our epistemic capacities to apply them. If Soames rejects V as well, he must do more to explain what is gained by rejecting bivalence for vague predicates.

In addition, throughout the chapter Soames seems to confuse two issues: (1) Do the semantic rules for vague predicates (at a given point of time, in a fixed context) support bivalence? (2) Do the semantic rules for vague predicates shift dynamically as the context changes? An epistemicist could easily answer 'No' to the first question and 'Yes' to the second, thereby gaining all the advantages of a dynamical account while avoiding the cost of truth-value gaps.

8. Deflationary Theories

In the final chapter Soames discusses deflationary theories of truth. He surveys a variety of deflationist approaches including Ramsey's redundancy theory, Strawson's performative theory, Tarski's and Kripke's theories of truth (construed as deflationary), and Horwich's minimalism. Soames makes a distinction between those concepts that are *philosophically contentious* and *substantive* and those that are not. Deflationism consists in the view that *truth* belongs in the second category, being adequately explicated by the Tarski biconditionals.

Since Soames accepts that propositions exist and are truth-bearers, he concedes that uses of 'true' are genuine predications. Hence, he concedes that there must exist a *property* of truth. In this concession, however, Soames uses the concept of *property* in a minimalist way. In contrast, many modern-day Platonists (such as Armstrong, Tooley, and myself) reserve the concept of *property* for the making of a more substantial ontological claim. Platonists do not accept a principle of comprehension for properties: they do not accept that a property exists corresponding to every meaningful predication.

In addition, Soames does not mention any substantive theory of properties such as the Frege structures of Aczel or Turner, nor does he refer to work on properties from the perspective of relevance logic such as that of Dunn. Hence, Soames's concession that a property of truth exists must be construed as ontologically minimal.

In discussing Horwich's minimalist theory, Soames relies on an objection also raised by Gupta in [8]. This objection consists of pointing out that Horwich's theory cannot explain many important generalizations about truth including the following:

(26) For any propositions p and q, the conjunction of p and q is true iff p is true and q is true.

Horwich's theory can explain every *instance* of (26), but (26) itself is not a logical consequence of Horwich's theory, as Gupta also pointed out. The set of instances of (26) do not entail (26) without the addition of the piece of information that the set contains instances for every (relevant) proposition.

Finally, Soames considers a recent argument by Boghossian to the effect that deflationism about truth is self-refuting. Boghossian's argument consists of three theses:

Thesis 1 Deflationism about truth is incompatible with nonfactualism.

Thesis 2 Nonfactualism about psychological or linguistic content entails deflationism and so is inconsistent.

Thesis 3 Deflationism about truth is a form of nonfactualism and so is inconsistent.

Boghossian takes deflationism about truth to be the thesis that there is no property of truth. He defines nonfactualism about a discourse X as the thesis that no sentence constructible within X has the property of truth. Thesis 1 follows from these definitions since nonfactualism of any kind implies the existence of the property of truth since it denies that that very property of truth is not possessed by some set of sentences. Thesis 3 is supposed to follow since, if there is no property of truth, sentences apparently predicating truth are not genuine assertions and so fail to have the property of truth.

The claim that there is no property of truth can itself be taken in two ways: as an ontological claim, or as a logical or semantical one. On the first reading, the denial of the existence of a *property* of truth does not entail that predications of truth fail to be genuine predications at all. As I mentioned above, Platonists typically take a claim of the existence of a property as involving more than just a claim about the logical form of certain predications: it is, in addition, a claim about the causal or explanatory structure of the world (see Lewis [13]). Platonists have often denied that certain logical predications, including the predication of truth, involve any such ontological commitment.

Construed ontologically, deflationism is certainly compatible with nonfactualism about any universe of discourse. It is only when deflationism is construed as making a claim about the logical form of predications of truth that Thesis 1 is plausible. Soames comes to a similar conclusion. He notes that Boghossian assumes something like the following:

(B) To hold that a predicate F does not express a property is to be a nonfactualist about discourse involving F.

If the existence of a property is construed as a substantive ontological matter, then (B) clearly fails to hold.

As Soames points out, Boghossian's Thesis 2 depends on Thesis 1, and so it too depends on construing deflationism in logical, non-ontological terms. In addition, Boghossian's argument for Thesis 2 depends on taking deflationism to be a thesis about the truth of *sentences*, not of *propositions*. One could, as Soames in fact does, hold a strongly deflationist position on the truth of propositions while accepting that there is, logically speaking, a genuine property of truth for sentences, namely, the property of expressing a true proposition. Hence, Boghossian's Thesis 2 is false under two construals of *deflationism*: (1) if deflationism is a relatively weak, ontological claim, and not a logical one, and (2) if deflationism consists in a claim (even a strong, logical claim) about the truth of propositions only.

References

[1] Barwise, J., and J. Etchemendy, *The Liar: An Essay on Truth and Circularity*, Clarendon Press, New York, 1987. Zbl 0678.03001. MR 88k:03009.

[2] Brandom, R., *Making It Explicit: Reasoning, Representing, and Discursive Commitment*, Harvard University Press, Cambridge, 1994.

[3] Burge, T., "Semantical paradox," *Journal of Philosophy*, vol. 76 (1979), pp. 169–98.

[4] Etchemendy, J., *The Concept of Logical Consequence*, Harvard University Press, Cambridge, 1990. Zbl 0743.03002. MR 92f:03003.

[5] Gaifman, H., "Pointers to truth," *The Journal of Philosophy*, vol. 89 (1992), pp. 223–61. MR 93h:03002.

[6] Glanzberg, M., "The liar in context," *Philosophical Studies*, vol. 103 (2001), pp. 217–51. MR 2002c:03007.

[7] Gomez-Torrente, M., *Tarski's Definition of Logical Consequence: Historical and Philosophical Aspects*, PhD thesis, Princeton University, 1996.

[8] Gupta, A., "A critique of deflationism," *Philosophical Topics*, vol. 21 (1993), pp. 57–81.

[9] Gupta, A., and N. Belnap, *The Revision Theory of Truth*, The MIT Press, Cambridge, 1993. Zbl 0858.03010. MR 95f:03003.

[10] Koons, R. C., *Paradoxes of Belief and Strategic Rationality*, Cambridge University Press, Cambridge, 1992. MR 93d:03029.

[11] Koons, R. C., "Circularity and hierarchy," pp. 177–98 in *Circularity, Definition and Truth*, edited by A. Chapuis and A. Gupta, Indian Council of Philosophical Research, New Delhi, 2000.

[12] Kripke, S., "Outline of a theory of truth," *Journal of Philosophy*, vol. 72 (1975), pp. 690–716. Zbl 0952.03513.

[13] Lewis, D., "New work for a theory of universals," *Australasian Journal of Philosophy*, vol. 61 (1983), pp. 343–77.

[14] McGee, V., *Truth, Vagueness and Paradox: An Essay on the Logic of Truth*, Hackett Publishing, Indianapolis, 1990. . Zbl 0734.03001. MR 92k:03004.

[15] Parsons, T., "Assertion, denial, and the Liar Paradox," *Journal of Philosophical Logic*, vol. 13 (1984), pp. 137–52. MR 86g:03009.

[16] Platts, M., *Ways of Meaning: An Introduction to a Philosophy of Language*, Routledge and Kegan Paul, London, 1979.

[17] Ray, G., "Logical consequence: a defense of Tarski," *Journal of Philosophical Logic*, vol. 25 (1996), pp. 617–77. Zbl 0870.03005. MR 98a:03009.

[18] Reichenbach, H., *Experience and Prediction: An Analysis of the Foundations and Structure of Knowledge*, University of Chicago Press, Chicago, 1961.

[19] Simmons, K., "The diagonal argument and the liar," *Journal of Philosophical Logic*, vol. 19 (1990), pp. 277–303. Zbl 0702.03001. MR 91k:03013.

[20] Simmons, K., *Universality and the Liar: An Essay on Truth and the Diagonal Argument*, Cambridge University Press, Cambridge, 1993. Zbl 0977.03502. MR 94m:03017.

[21] Sorensen, R., *Blindspots*, Clarendon Press, Oxford, 1988.

[22] Tarski, A., "On the concept of logical consequence," pp. 97–112 in *Logic, Semantics, Metamathematics*, Clarendon Press, Oxford, 2d edition, 1956. Papers from 1923 to 1938, translated by J. H. Woodger, edited and with an introduction by John Corcoran. Zbl 0075.00702. MR 17,1171a.

[23] Williamson, T., *Vagueness*, Routledge, London, 1994.

Robert C. Koons
Department of Philosophy
University of Texas at Austin
Waggener Hall 405
Austin TX 78712-1180
rkoons@mail.utexas.edu
http://www.utexas.edu/cola/depts/philosophy/faculty/koons/main.html