Tarski and Primitivism About Truth

Jamin Asay

Lingnan University

1. Introduction

The view that truth is a primitive, indefinable notion was of central importance to the originators of analytic philosophy. Moore and Russell adopted the view after abandoning their idealism (though they soon turned to correspondence accounts), and Frege subscribed to it until the end of his life.[1] But save for some attention given to the view by Davidson (1990, 1996), primitivism about truth has laid low for the last century. During that time, by far the dominant force in the theory of truth has been Tarski, and much subsequent discussion has been focused around the question of whether Tarski's work better motivates a robust, correspondence-style theory of truth,[2] or a more deflationary approach.[3] I reject this dichotomy, and argue in this paper that Tarski's work on truth is actually most consonant with a primitivist perspective on truth; hence, his views should not be thought to lend support to either correspondence or deflationary theories. Given that Tarski shows how to offer a definition of truth, the congeniality between his views and primitivism may not be immediately obvious, and my aim is to draw the appropriate connections. I do not argue that Tarski himself subscribed to a primitivist conception of truth, though I shall show how the view is open to him, and is more amenable to his views on truth than are the more familiar theories of truth.

1. See, *inter alia*, Moore 1899, Russell 1904, and Frege 1956.

2. See, *e.g.*, Popper 1963: 223–228 and 1979, Davidson 1969: 758, Jennings 1987, Horwich 1990: 9, Kirkham 1992: 170, Niiniluoto 1999, and Fernández Moreno 2001. Woleński argues that Tarski employs a form of "weak correspondence" (1993). Others demur from the correspondence interpretation: see, *e.g.*, Haack 1976 and 1978, Keuth 1978, Putnam 1985: 72, Davidson 1990, 1996: 268, and 1999: 110, Künne 2003: 213, and Simmons 2009: 555–556.

3. See, *e.g.*, Black 1948: 63, Horwich 1982, Soames 1984: 416 and 1999: 238, Stoutland 1999: 83, and Patterson 2010: 25. Those who argue against the deflationary interpretation include Davidson 1990, 1996: 269, and 1999: 110, Schantz 1998, Ketland 1999, Horwich 2005, and Simmons 2009: 556–558. See Field 1994: 269–270 for an intermediate view. Obviously, much turns on what exactly is meant by 'correspondence theory' and 'deflationism', and so some subtlety in interpreting these attributions is called for. For general discussion about how to interpret Tarski's view *vis-à-vis* correspondence and deflationism, see Kirkham 1992: 170–173, Künne 2003: 208–213, and Patterson 2012: 140–143.

To begin, I explain what primitivism is, and how it offers a perspective on truth that competes with both correspondence and deflationary theories. I then review the aspects of Tarski's work on truth that are most relevant to the primitivist theory of truth. In so doing, I show how Tarski is a natural ally to defenders of primitivism, and how primitivism is a favorable view for Tarski to hold. I conclude by showing how we can use primitivism to help frame and respond to some familiar charges, due to Putnam and Etchemendy, that have been brought against the merits of Tarski's project.

2. Primitivism About Truth

The primitivist theory of truth that I shall be addressing offers a *substantive* account of the *concept* of truth. The theory is substantive (*i.e.*, non-deflationary) in that it admits that truth is a philosophically important notion, one that has explanatory value that outstrips its linguistic and expressive features. By contrast, deflationary theories of truth typically hold that all there is to the theory of truth is an account of truth's utility in disquotation, forming generalizations, expressing infinite conjunctions and disjunctions, etc. Hence, for deflationists, truth has no *explanatory* role to play in philosophy; the truth predicate is merely of *expressive* use in giving accounts of other notions such as assertion, belief, and meaning.[4]

Furthermore, the primitivism I am defending involves most fundamentally the *concept* of truth, as opposed to the *property* of truth (or *words* like 'true' and 'truth'). Whether or not the property of truth should be understood in a "metaphysically robust" way is a separate question, and one that I shall set aside for present purposes.[5] Discussion of the property of truth is a fundamentally metaphysical enterprise that concerns the feature(s) that truth bearers possess when they're true, and

what it is that makes truth bearers true; discussion of the concept of truth, by contrast, focuses on what it is for us as cognitive agents to possess a notion of truth, and what it means for us to deploy it in our thought and language.

Primitivism is the view that truth is a fundamental concept. As such, it cannot be analyzed, defined, or reduced into concepts that are more fundamental.[6] Accordingly, primitivism is opposed to the traditional substantive theories of truth, such as the correspondence, coherence, and pragmatic theories, which hold that truth, while a substantive notion, is nevertheless analyzable into those further notions. Primitivism is also opposed to deflationism, which rejects the claim of truth's fundamentality. Horwich, for example, argues that truth can be defined, albeit implicitly, by the T-sentences (*e.g.*, 'The proposition that 2 is prime is true if and only if 2 is prime'), which he claims are the "fundamental principles" of truth (1990: 18). Primitivists, by contrast, deny that there are any such fundamental principles about truth. If there were, truth itself wouldn't be fundamental. Far from being a fundamental concept of preeminent philosophical worth, truth, say deflationists, is a rather innocuous notion of mere expressive utility.

Primitivism, like the other theories of truth, aims to offer an account of our most general concept of truth, whatever it is that is shared by all who can be said to possess the concept, regardless of which language they speak or how complex their thoughts may be. Tarski never attempts to give such an account, and at times makes remarks that might appear to disparage any such endeavor. He observes, for example, that no one has adequately explained to him what "the philosophical problem of truth" is, and that he does not understand what it is to offer the "essence" of a concept like truth (1944: 361). The primitivist acknowledges that truth has no essence, at least if something's essence is constituted by its *analysans*. Primitivism takes as its target our most basic notion of truth, against which the adequacy of

---

4. See, *e.g.*, chapter 3 of Horwich 1990, Field 1994, and Williams 1999.

5. In my forthcoming, I defend a *metaphysically deflationist* account of the property of truth: I argue that truth is best understood merely as an *abundant* property, and not a *sparse* property, regardless of how one might draw that distinction. See Lewis 1983 for an account of sparse and abundant properties.

6. Some primitivists admit that truth can be defined, albeit in terms of other concepts that are equally fundamental (thereby forming a circle of interdefinable fundamental concepts). See Strawson 1992 for a view of this sort.

Tarski's definitions are tested. Tarski is rightly skeptical that that basic notion admits of any precise definition; but that does not mean that he thinks there is no such notion. As we shall see, Tarski thinks that we do have an intuitive conception of truth (however vague or ill-defined) that is prior to how it is employed in language. That conception is the primitivist's focus.

For our purposes, the dimensions of primitivism most relevant to Tarski's work on truth are truth's indefinability, explanatory value, meta-physically neutral character, and relationship to the T-sentences. These aspects of the primitivist's conception of truth are shared with Tarski's basic perspective on truth, and will form the basis of our discussion.

3. Tarski's Theory of Truth

In his seminal work on truth, Tarski (1944, 1956a) advances and defends what he calls the "semantic conception" of truth. His main goal is to offer a satisfactory definition of truth for a language that (i) avoids paradox, (ii) deflects positivist skepticism, and (iii) remains true to our "classical conception" of truth.

Consider first the third desideratum. "The desired definition," Tarski writes, "does not aim to specify the meaning of a familiar word used to denote a novel notion; on the contrary, it aims to catch hold of the actual meaning of an old notion" (1944: 341). The old notion is the intuitive view of truth that Tarski attributes to Aristotle, and which he thinks contemporary correspondence theorists are attempting to uphold. According to Aristotle, "A falsity is a statement of that which is that it is not, or of that which is not that it is; and a truth is a statement of that which is that it is, or of that which is not that it is not" (1966: 70; 1011b25–28). Tarski saw his work as offering "a more precise expression of our intuitions" to the effect that truth somehow consists in a kind of correspondence with reality (1944: 343; cf. 1956a: 153). Hence, it's understandable why many have taken Tarski's work to motivate — or even just be — a kind of correspondence theory of truth.

So Tarski clearly thinks that his work aims to do justice to our classical conception of truth, a goal that he takes himself to share with correspondence theorists (though not pragmatists or coherentists). In fact, Tarski thinks his work outperforms correspondence theory: "I do not have any doubts that our formulation does conform to the intuitive content of that of Aristotle. I am less certain regarding the latter [correspondence] formulations of the classical conception, for they are very vague indeed" (1944: 360).

By offering a precise definition of truth, Tarski is hoping to make the notion safe from those who have voiced skepticism about the very idea of truth. Semantic concepts like truth, Tarski writes,

> have been treated for a long time with a certain amount of suspicion. From a historical standpoint, this suspicion is to be regarded as completely justified. For although the meaning of semantic concepts as they are used in every-day language seems to be rather clear and understandable, still all attempts to characterize this meaning in a general and exact way miscarried. (1944: 346; cf. 1956b: 401)

Bearing in mind that Tarski's work on truth was developed in the hey-day of logical positivism, his concern here is understandable. His intent is to offer a precise definition of truth that is free of any "alleged metaphysical implications", so that any lingering distrust in the notion of truth may "evaporate" (1944: 364; cf. 1956a: 252).

Part of the project of making the notion of truth "safe" for philosophy and science involves developing a metaphysically neutral account of the concept; the other crucial part is showing that the notion is not infected with paradox. The semantic paradoxes — including, most notoriously, the liar paradox — threaten to undermine the legitimacy of the notion of truth. To see why, consider that in order to ensure that the semantic conception of truth conforms to the classical Aristotelian conception, Tarski offers his famous adequacy condition. In order for a given language's definition of truth to be adequate it must, for every sentence of the language, imply an instance of the schema:

(T)   *X* is true if and only if *p*.

The '*X*' is to be replaced by a *name* of a sentence of the target language whose truth predicate is being defined, and '*p*' by a translation of that sentence in the language doing the defining.[7] A definition that meets this condition offers necessary and sufficient conditions for the truth of every sentence of the language in question in a way that fits with Aristotle's dictum (1944: 344). The adequacy condition reveals that Tarski thinks we do have a pre-linguistic conception of truth, and that any purported definition of truth had better conform to it.[8]

Paradox arises when we apply sentences like the following to the truth schema:

(L)   The first named sentence in section 3 of "Tarski and Primitivism About Truth" is not true.

If we plug (L) into our schema, we have:

(T_L) (L) is true if and only if the first named sentence in section 3 of "Tarski and Primitivism About Truth" is not true.

Bearing in mind the identity of (L), we can see that (T_L) is equivalent to:

(T_L*)(L) is true if and only if (L) is not true.

(T_L*) is contradictory, and so any language that allows the formation of (L) and meets the adequacy condition concerning schema (T) will produce a contradiction (assuming classical logic).

Tarski's positive account of truth is formulated so as to be inoculated from paradox. Most crucially, Tarski offers definitions of truth that are *relative to a particular language* (1944: 342). Rather than offering a general definition of truth that applies across all languages, Tarski defines language-relative truth predicates, such as 'true-in-L_1', 'true-in-L_2', and the like. Secondly, truth cannot be defined *within* the language for which it is being defined. Instead, truth must be defined in the *metalanguage*. The metalanguage for an *object language* is a language

that includes names for all the sentences of the object language and other resources necessary for defining semantic predicates for the object language. If truth is to be defined for some language O, then O cannot include the truth predicate 'true-in-O'. The predicate is instead a part of the metalanguage M. Because O does not contain its own truth predicate, liar sentences cannot be formulated within O, and so paradox is avoided.

Hence, for Tarski, truth is definable only relative to certain sorts of languages. Those languages cannot be "semantically closed", which is to say that they cannot include their own semantic machinery, such as names for their sentences and their own truth predicate. Truth predicates for a given object language can be defined only from within a metalanguage that is expressively more powerful than the object language.

Given these constraints, we can now offer a definition of 'true-in-O' for any qualified object language O. The definition is constructed via the notion of satisfaction. Satisfaction, ordinarily, is a relation between objects and the sentential functions that constitute a given language. First we define the satisfaction relation by listing out the most basic cases: for example, Socrates satisfies the sentential function '*x* is a philosopher', and the sequence <Socrates, Plato> satisfies the sentential function '*x* is a teacher of *y*'. We then define satisfaction recursively by showing how the satisfaction conditions for compound sentential functions (and those involving quantification) are to be given in terms of the satisfaction conditions for the basic functions.

For Tarski, the definition of satisfaction is somewhat more involved, since he defines it as a relation between sentential functions and infinite sequences of objects. To see how the definition works, we first need to establish a correlation between variables and the objects in the infinite sequences. If we set up the list of variables in the standard way (*i.e.*, <*a, b, c, d, …, x, y, z, a_1, …*>), then *x* is correlated with the twenty-fourth object in the sequence. Now we can say that a sequence satisfies '*x* is a philosopher' if and only if the twenty-fourth object in the sequence is such that it is a philosopher. A sequence satisfies '*x* is a

---

7.   Whether Tarski intends the condition as a necessary or sufficient one is contentious; see Patterson 2006a for discussion.

8.   Cf. Heck 1997: 541–542.

teacher of *y*' just in case the twenty-fourth member of the sequence is a teacher of the twenty-fifth member of the sequence. Now consider sentences, which are just sentential functions with no free variables. Any sequence whatsoever satisfies 'Socrates is a philosopher' and 'Socrates is a teacher of Plato', as they involve no variables at all. For the existentially quantified sentence 'There is some *x* such that *x* is a philosopher', a sequence S satisfies it just in case there is a sequence S' that differs from S at most with respect to the objects in their twenty-fourth positions, and whose twenty-fourth object is a philosopher. So long as there is some philosopher in existence, there will be an S' with a philosopher in the twenty-fourth position. As a result, the existential will be satisfied by every infinite sequence should there be a philosopher, and satisfied by no infinite sequence should there not be any philosophers.

This procedure defines the satisfaction relation for a given language, which can now be used to define truth itself. Sentences are true-in-O if and only if they are satisfied$_O$ by each and every infinite sequence of objects; sentences are false-in-O if and only if they are satisfied$_O$ by no object or sequence of objects (1944: 353; 1956a: 195). Hence we have arrived at Tarski's definition of truth: true sentences are those that are satisfied by every infinite sequence of objects.

4. Tarski and Primitivism

We have now seen Tarski's semantic conception of truth in outline, and the method he provides for defining truth for a given language. My contentions are that, of all the contemporary theories of truth, primitivism is the most compatible with the semantic conception, and that primitivists can take advantage of Tarski's work for their own purposes. Traditionally, Tarski's work has been thought to motivate either correspondence accounts or deflationary accounts; I reject both suggestions. To see why, let us turn now to four crucial aspects of Tarski's work, namely, his contention that truth is ultimately indefinable, his admission that truth is explanatorily valuable, the relationship

between truth and the T-sentences, and the metaphysically neutral character of truth.

4.1. Indefinability

Primitivism begins with the thesis that truth cannot be defined or analyzed. So how can Tarski's work, which shows how to offer precise definitions of truth, motivate primitivism and its key commitment to the indefinability of truth? To resolve this tension, we need only attend to the limitations inherent to Tarski's method. As we have seen, Tarski shows how to define truth *for a particular language*, and only for *languages of a particular type*. The languages in question are (i) *formal* languages which are (ii) not *semantically closed* but (iii) satisfy the condition of *essential richness*. Absent these features, truth is indefinable.

First, the language needs to be *formal* (1956a: 165). This criterion is important because it enables us to offer a precise definition of the notion of satisfaction, which relies on there being a precise account of what the sentential functions of the language are. Tarski writes: *"The problem of the definition of truth obtains a precise meaning and can be solved in a rigorous way only for those languages whose structure has been exactly specified*. For other languages — thus, for all natural, "spoken" languages — the meaning of the problem is more or less vague, and its solution can have only an approximate character" (1944: 347). "Our everyday language," Tarski goes on to say, "is certainly not one with an exactly specified structure. We do not know precisely which expressions are sentences" (1944: 349). Tarski thus draws a crucial distinction between natural and formal languages, and is explicit that his methods apply only to the latter.

Next, Tarski requires that the languages for which truth can be defined not be "semantically closed" (1944: 348). A semantically closed language is, essentially, a language that includes its own semantics. Hence, a semantically closed language includes names for all its constitutive sentences; if the sentence 'Snow is white' belongs to the language, so too does the name of that sentence, ''Snow is white''.

Semantically closed languages also include their own truth predicate, whose extension includes sentences of that very language. As a result, semantically closed languages enable the formation of paradoxical liar sentences, which disqualifies them from Tarski's method. Natural languages exhibit semantic closure by way of their "universality": "A characteristic feature of colloquial language (in contrast to various scientific languages) is its universality. It would not be in harmony with the spirit of this language if in some other language a word occurred which could not be translated into it" (1956a: 164). Hence, in his positive account, Tarski must rely on the object language/metalanguage distinction, for the truth predicate to be defined cannot belong to the language to which it applies.

One final criterion remains: the metalanguage defining truth for its object language must satisfy the condition of "essential richness". What is required is that the metalanguage doing the defining be essentially richer than the object language. Tarski notes that it is not easy to give a precise characterization of essential richness.[9] What he does say is that "If we restrict ourselves to languages based on the logical theory of types, the condition for the meta-language to be "essentially richer" than the object-language is that it contain variables of a higher logical type than those of the object-language" (1944: 352). If the metalanguage fails to be richer than the object language, then an interpretation of the metalanguage can be offered within the object language. This, in turn, will permit the possibility of constructing self-referential sentences that enable the semantic antinomies to resurface. Hence, the condition of essential richness is in place in order to ensure that the right sort of relationship exists between the object language and the metalanguage. The metalanguage needs to be richer than the object language so that it can offer a "broader" or "higher" perspective from which to contain and define truth in the object language. If the metalanguage is on the same level as the object

language, so to speak, then the semantic paradoxes will be at play, and the definition won't succeed.[10]

In fact, Tarski offers a proof for why truth cannot be defined within a language (or from a metalanguage of the same or lower order). Tarski's formal "indefinability" proof employs some of the same diagonal reasoning that lies behind Gödel's work on incompleteness. I shall offer only a brief informal sketch of the proof here, which basically presents the proof as a form of Grelling's heterological paradox (a connection Tarski observes at his 1956a: 248, footnote 2).[11] For some language of sufficient expressive power L, we first suppose that there is a predicate 'Tr' such that its extension contains all and only true sentences of L. But we can now form a new predicate, 'is not Tr of itself', and ask whether it holds of itself or not. If 'is not Tr of itself' is not Tr of itself, then it belongs in its own extension, and so is Tr of itself after all. But if 'is not Tr of itself' is Tr of itself, then it is not Tr of itself, given that the predicate holds only of predicates that are not Tr of themselves. Either way, contradiction follows. The assumption reduced to absurdity is that there is such a predicate 'Tr' that contains all and only true sentences of L. There is no such predicate, and so truth is not definable for L.

10. In the original version of "The Concept of Truth in Formalized Languages", the languages that satisfy the condition of essential richness are the "languages of finite order" discussed in section 4, such as the language of the calculus of classes (1956a: 209), and simple first-order languages. Languages that don't satisfy the condition are the languages of infinite order discussed in section 5, such as the general theory of classes. For these languages, the satisfaction relation cannot be defined by the method previously outlined, and so no Tarski-style definition of truth can be constructed for them (1956a: 244). However, Tarski later revises his view on this point in the postscript to "The Concept of Truth in Formalized Languages" (1956a: 268–278). There, Tarski argues that truth in languages of infinite order *can* be defined, so long as it is defined from a language of a higher order, which is now possible given Tarski's embrace of the transfinite. Hence, what matters for Tarski is the nature of the relationship between the object language and metalanguage, and not whether the object language in question is of finite or infinite order.

11. For more formal presentations of Tarski's proof, see Simmons 1990: 288–289, 1993, and 2009: 550–553, Patterson 2006a: 16–24 and 2012: 144–160, and chapter 2 of Field 2008. See Martin 1968 for more on Grelling's paradox.

9. See DeVidi and Solomon 1999 for an argument that no tenable account of essential richness is available to Tarski, and Ray 2005 for a defense of Tarski.

Where the condition of essential richness is not satisfied, truth cannot be defined in Tarskian fashion. Interestingly, however, Tarski notes that a different road is left open to these languages. Though truth cannot be defined within them, we can introduce a *primitive* term 'true' and then give an account of it, not by way of definition, but axiomatization:

> If we want to develop the theory of truth in a meta-language which does not satisfy this condition [of essential richness], we must give up the idea of defining truth with the exclusive help of those terms which were indicated above […]. We have then to include the term "*true*," or some other semantic term, in the list of undefined terms of the meta-language, and to express fundamental properties of the notion of truth in a series of axioms. There is nothing essentially wrong in such an axiomatic procedure, and it may prove useful for various purposes. (1944: 352)

Tarski goes on to provide some of the details as to how such an axiomatization would proceed (1956a: 255–265). Here we see one way in which Tarski thinks we can make use of a primitive notion of truth. For some of the languages for which truth cannot be defined, we can nevertheless make "consistent and correct use" of the concept of truth by way of taking truth as a primitive notion, and giving it content by introducing the relevant sorts of axioms (1956a: 266).[12]

Let us now return the discussion to conceptual primitivism. The primitivism about truth that I am defending is not about truth as defined for any particular language. Primitivism concerns our *concept* of truth, which in turn gives content to the various truth predicates and operators that occur in our natural and formal languages. So we must be cautious about drawing implications for primitivism too quickly from Tarski's work. Nevertheless, a few points of connection are important to note.

12. See Halbach 2011 and Horsten 2011 for contemporary treatments of the axiomatic approach to truth.

First, Tarski despairs of the attempt to offer a definition of truth in natural language. If any truth predicate were to express our most general notion of truth (for which primitivism hopes to account), it would be the truth predicate of natural language. But Tarski is committed to the view that his methods cannot offer a definition of that most wide-ranging truth predicate, for natural languages fail his requirements of formality and of not being semantically closed As it turns out, Tarski thinks the matter goes beyond indefinability: given natural language's semantic universality, the very consistency of our use of truth within it is suspect.[13] Tarski writes: "the very possibility of a consistent use of the expression 'true sentence' which is in harmony with the laws of logic and the spirit of everyday language seems to be very questionable, and consequently the same doubt attaches to the possibility of constructing a correct definition of this expression" (1956a: 165; emphasis removed). Note Tarski's hedged language here. He does not claim to have "proved" that truth in natural language is indefinable or inconsistent.[14] Later he writes that "We may at best only risk the guess that a language whose structure has been exactly specified and which resembles our everyday language as closely as possible would be inconsistent" (1944: 349). Absent a definitive account of the nature, structure, and limits of natural language, Tarski refuses to make a definitive pronouncement on the definability of truth in natural language (1944: 347).

13. See Ray 2003 and Patterson 2006b and 2012: 160–168 for more on Tarski and the potential inconsistency of natural languages.

14. Davidson repeatedly claims that Tarski has "proved" that the concept of truth is indefinable (1990: 285–286, 1996: 265, 269–270, 275–276), but it's unclear what proof he is referring to. Tarski's formal indefinability proof (as found in section 5 of Tarski 1956a) applies only to formal languages that fail to satisfy the condition of essential richness. As for Tarski's pessimistic remarks concerning the definability of truth for natural languages (as found in section 1 of Tarski 1956a), these do not constitute a proof, as the hedged language of the passages quoted in this paragraph reveals. Like Davidson, I believe that primitivists can find plenty of value in Tarski's work on definability, but I do not believe, as Davidson appears to, that Tarski has proven the truth of primitivism (or that Tarski believes himself to have done so). See also García-Carpintero 1999: 142–143.

Despite Tarski's caution, primitivists may take some solace in his conclusions regarding natural language. Primitivists claim that our concept of truth admits of no analysis. If any purported definition of a truth predicate were to falsify that claim, it would be a definition of the truth predicate of natural language.[15] At the very least, Tarski shows that his methods for definition do not apply to the natural language 'true'. So the lesson primitivists may draw from Tarski is that one potential source of falsification to their view is of no threat. Hence, the primitivist may argue as follows: if, *contra* primitivism, our concept of truth is definable, then the natural language truth predicate should admit of a Tarski-like definition. However, no such definition is available, and thus our concept of truth is not definable. Despite the fact that various formal languages and fragments of natural language admit of a definable truth predicate, our most basic and general concept of truth allows no such definition. Of course, whether there are other, non-Tarskian ways of defining our most basic concept of truth is an open question, and Tarski is wise not to pronounce definitively on the possibility of other methods. Still, primitivists may share Tarski's suspicion that the prospects for definition are grim.

The primitivist argues that our *general*, *inter-linguistic* notion of truth is what cannot be defined. Tarski appears to be in full agreement with this particular thesis. His entire approach to the theory of truth can be taken as an argument in favor of it (though not, as we have seen, some kind of conclusive proof). Truth can be defined only with respect to certain limited kinds of languages. As for our most basic notion of truth, definition is not to be had. Hence, even though Tarski's definitions do

justice to our classical, intuitive conception of truth, they by no means *exhaust* that conception.

Tarski's work on truth, at the end of the day, offers no positive solution to the problem of defining the notion of truth at its most basic and fundamental level. As a result, Tarski would reject the traditional substantive theories of truth that intend to do precisely that. In particular, Tarski would reject the correspondence theorist's attempt to define truth in terms of the more fundamental notions of *correspondence* and *reality* (or *facts*). But as we saw, Tarski thought of himself as partly engaged in the same project as correspondence theorists, in the sense that both are interested in giving an account of truth that adheres to the traditional classical conception of truth. (Hence, as we have seen, many commentators have interpreted Tarski as a correspondence theorist.) What we may appreciate is that Tarski thinks that the best way of staying true to the classical conception is not by analyzing truth by way of correspondence, but by holding one's theory of truth accountable to the T-sentences formed from schema (T). As a result, we do not see Tarski try to give accounts of the notion of *correspondence* and *fact*, and then use those notions to define *truth*. Here we have another point of intersection between Tarski and primitivism: the best way to account for the "correspondence intuition" (*i. e.*, whatever basic understanding of truth it is that Aristotle's formulation intends to capture) is not by way of *defining* truth in terms of correspondence, but by way of adherence to the T-sentences.[16] Tarski, like the primitivist, rejects correspondence theory (understood as the project of trying to define truth in terms of correspondence) as a way of understanding the nature of truth itself. The correspondence theorist's definition is not necessary for capturing what's right about the thought that truth in some sense consists in a kind of correspondence with the world.

---

15. I happen to doubt the truth of the antecedent here, since I believe that our concept of truth is prior to its use in our language, and is not fully exhausted by the use that 'is true' plays in our language. (For example, it's not clear how a Tarski-style definition of truth can illuminate what it means to say *that to assert is to present as true*. Cf. Bar-On and Simmons 2007 and Textor 2010.) But I suspect that Davidson believes something along these lines, since he appears to take Tarski's negative results concerning the possibility of defining a fully general, inter-linguistic truth *predicate* to establish the primitive nature of our *concept* of truth.

16. Below, and in my 2011, I go further and argue that attention to *truthmaking*, and not the nature of truth itself, is vital to understanding what's right about the "correspondence intuition". The best way to capture the correspondence intuition, I argue, is not simply by upholding the T-sentences, but by offering in addition a theory of truthmaking.

Finally, it's worth noting that even for those languages where a Tarskian definition of truth is available, it's unclear that his definitions capture the real nature — or *intension*, we might say — of truth. If they don't, then Tarski's definitions do not in any way conflict with primitivists' claims regarding the indefinability of truth. For primitivists argue that no definition of truth can reveal the nature of truth in more fundamental terms. If Tarski isn't even intending to do that — see Patterson 2008a: 178 — then we can appreciate how Tarski's project is in full harmony with primitivists'. Tarski defines true sentences, where they are definable, as those that are satisfied by every infinite sequence of objects. As Tarski shows, this definition is "materially adequate" and "formally correct" (1944: 341, 1956a: 152). That is to say, it entails all the T-sentences (and therefore captures our intuitive notion of truth, and not something else), and suffers no internal problems such as circularity. Furthermore, the definition captures the right extension: it includes all and only true sentences of the language. However, it's quite implausible to think that Tarski's definitions have hit on the correct intension or meaning of truth.[17] One way to press the point is as follows. Granted, sentences are true if and only if they are satisfied by all infinite sequences of objects. But are they true *because* they are satisfied by each and every infinite sequence of objects? Is the right account of the truth of 'Snow is white' that every infinite sequence satisfies it, or does the right account have something to do in particular with snow and its color? If the former answer seems insufficient (as it does to me), then there is reason to suppose that Tarski's definitions were never intended to capture the nature of truth and its intension in more fundamental or basic terms.[18]

This conclusion becomes all the more apparent given Tarski's recognition that for some languages, truth is not to be cashed out in terms of satisfaction at all, but rather by a process of axiomatization. So we may conclude that, formally speaking, truth is not "univocal" as it appears in formal languages that do or do not satisfy the condition of essential richness. Sometimes truth is defined in terms of satisfaction, sometimes it is introduced through axiomatization. But if we believe there to be one more basic notion at hand behind both kinds of languages, then we must reject the idea that truth is to be understood in terms of the more fundamental notion of satisfaction. That's not the kind of analysis that Tarski offers. Such an analysis *is* what correspondence theorists, coherence theorists, and pragmatists offer. They define truth in terms of more fundamental notions. But Tarski is not engaged in such a process, even when he is offering definitions of truth. Consequently, Tarski's work on truth leaves it entirely open how we are to understand the basic meaning or intension of truth. The most natural position for Tarski to take is that we cannot define this more general notion of truth; here again we see Tarski and primitivism being natural allies.

All told, where Tarski shows truth to be indefinable, primitivists can readily agree. The kind of indefinability of truth that primitivists advocate is precisely the kind of indefinability that Tarski's work allows for and motivates. Where Tarski advocates the definability of truth, it's clear that his definitions should not be thought of as capturing the nature of truth in more fundamental terms. That is to say, while his definitions represent an important kind of formal achievement, they should not be thought of as capturing the nature of truth in a way similar to the attempts of correspondence, coherence, and pragmatic theorists. Nor is there any reason to think that this result caused Tarski any despair. As we shall see in the next section, Tarski puts his definition to work in the service of a variety of mathematical and logical tasks (among others). Tarski's definitions of truth are best appreciated in light of these more formal accomplishments.

17. For one thing, Tarski's discussion of indefinability in section 5 of Tarski 1956a reveals that he thinks his satisfaction-based method of definition is in a certain sense arbitrary. Appealing to satisfaction for languages of finite order *works*, so he uses it. But Tarski is open to the idea that had satisfaction not worked, some other notion might have.

18. Cf. Patterson 2006b: 162–163.

4.2. *Explanatory value*

The previous section serves to illustrate some of the grounds Tarski has for rejecting the traditional substantive theories of truth (grounds that he may happily share with primitivists). Those theories purport to analyze in full our concept of truth, a project that Tarski rejects and may well deem impossible. In this section we can turn to why Tarski can also side with primitivists (and now against deflationists) in favor of truth's explanatory capabilities.

The first point to notice is that Tarski explicitly rejects the idea that if truth were to be definable and so in principle eliminable from language without expressive loss, then the notion of truth would thereby be "sterile":

> Consequently, the term "*true*" when occurring in a simple sentence of the form "*X is true*" can easily be eliminated, and the sentence itself, which belongs to the meta-language, can be replaced by an equivalent sentence of the object-language […]. Some people have therefore urged that the term "*true*" in the semantic sense can always be eliminated, and that for this reason the semantic conception of truth is altogether sterile and useless. And since the same considerations apply to other semantic notions, the conclusion has been drawn that semantics as a whole is a purely verbal game and at best only a harmless hobby. (1944: 358)

Not all deflationists would draw the conclusion that truth is a sterile and useless concept, but the standard deflationist line is that careful attention to truth's logical and linguistic role does reveal truth to carry little philosophical or explanatory worth (*e.g.*, Horwich 1990 and Williams 1999). But Tarski detects a fallacy in inferring sterility from definability:

> If, however, anyone continues to urge that—because of the theoretical possibility of eliminating the word "*true*"

on the basis of its definition—the concept of truth is sterile, he must accept the further conclusion that all defined notions are sterile. But this outcome is so absurd and so unsound historically that any comment on it is unnecessary. In fact, I am rather inclined to agree with those who maintain that the moments of greatest creative advancement in science frequently coincide with the introduction of new notions by means of definition. (1944: 359)

Hence, Tarski resists any suggestion that truth's importance is in any way compromised by establishing its definability or subsequent eliminability from the language. In fact, truth won't even belong to those languages for which truth can be defined! Nonetheless, understanding the notion of truth for a language is paramount for understanding that language.

For Tarski, the importance of seeking a definition of truth is not so as to eliminate the notion, but to put it to work. What can truth do for us? Tarski is adamant that his work on truth and semantics "can find applications in various domains of intellectual activity" (1944: 364). Tarski argues that his work can produce important results, not only for philosophy, but for both the empirical and deductive sciences.[19] As to philosophy, Tarski notes that his definition can be taken to have addressed "one of the fundamental problems of the theory of knowledge" (1956b: 407). Tarski also thinks that his approach to truth plays a vital role in "establishing semantics on a scientific basis" (*ibid.*), and thereby produces various empirical and logical fruits. His research has obvious direct ramifications for the study of linguistics, and "indirect" implications for the natural sciences, by way of the use of truth and other semantic notions within the methodology of science and scientific theory construction and acceptance (1944: 366–368). Furthermore, Tarski points out that his definition of truth leads to important results in logic, mathematics, and metamathematics (1944: 368–369). His definition of truth can be

19. See also Soames 1999: 100–107.

used to prove the metalogical versions of the laws of non-contradiction and of excluded middle (1956a: 197; cf. 1944: 354). It can also illuminate the notions of provability, definability, consistency, and completeness (1944: 368–369). In summarizing the potential positive benefits of his approach to truth for the deductive sciences, Tarski writes that "the establishment of scientific semantics, and in particular the definition of truth, enables us to match the negative results in the field of metamathematics with corresponding positive ones, and in that way to fill to some extent the gaps which have been revealed in the deductive method and in the very structure of deductive science" (1956b: 408).

In short, Tarski believes that the notion of truth can play an important, explanatory role in other areas of thought, inside and outside of philosophy. Rather than thinking of his project as having deflationary results, Tarski believed his work to be quite constructive, and only increase the import of the notion of truth. This attitude toward the theory of truth is shared by primitivism. For primitivists, truth is a foundational notion; its importance is demonstrated by showing how the notion can be put to work in other kinds of intellectual projects (including those that Tarski addresses). One of the reasons to take some notion as primitive is to show how it can then be put to various kinds of explanatory work. Primitivists may grant to deflationists that 'true' plays various kinds of logical and expressive roles in ordinary language, and that in many cases expressions with 'true' are equivalent to expressions that lack it. But primitivists and Tarski are in agreement that such facts about truth do not exhaust all that there is to be said about it; on the contrary, truth is a central, key notion that is highly relevant to our understanding of other important notions in science, mathematics, and philosophy.

4.3. *T-sentences*

A third connection between primitivism and Tarski (and one that also serves to further distance him from deflationism) involves the relationship between the T-sentences and the concept of truth. A standard deflationary interpretation of the T-sentences is that they are the

most fundamental principles about truth, in terms of which we grasp the concept (*e.g.*, Horwich 1990).[20] T-sentences thus play the role that other principles (which might involve correspondence, coherence, or utility) play in other theories. Primitivists reject the idea that there are any fundamental principles about truth, in terms of which the concept can be defined, reduced, or analyzed, and so must reject this deflationary interpretation of the T-sentences. On this point, Tarski can also be found to side with primitivists.

The T-sentences figure into Tarski's account of truth by way of serving as a kind of "check" on the material adequacy of any purported definition of truth; theories meet the condition when they entail all of the T-sentences. The notion of material adequacy at hand is not simply extensional adequacy (Patterson 2012: 109–111). Tarski's interest is instead in defining relativized truth predicates that conform to our intuitive notion of truth. Entailing the T-sentences is one way of guaranteeing that the notion defined hasn't strayed away from our most basic notion of truth. As Davidson puts the point, the T-sentences "alone constitute an unmistakable test that a theory has captured a concept of truth we are interested in" (1973: 77).

If the T-sentences are taken to be the fundamental facts about truth, in terms of which truth may be defined, then Tarski's adequacy condition is empty. If the deflationary conception of truth is defined by the T-sentences, then of course the deflationary conception will meet the adequacy condition by entailing them. Hence, the deflationary perspective on the truth schema and its T-sentences makes Tarski's adequacy condition vacuous and trivial. Since Tarski did not take the condition to be empty, we may appreciate at once how Tarski rejects the idea that the T-sentences provide the most fundamental facts about truth.[21] Primitivists, like Tarski, maintain that the T-sentences

20. See also the deflationary definitions of truth that involve the truth schema, as found (but not always endorsed) in Leeds 1978: 121, Field 1986: 58, Resnik 1990: 412, and David 1994: 107.

21. I hasten to add that this fact also shows that Tarski does not think of the T-sentences as being empirically empty logical truths, *contra* Putnam 1985, Etchemendy 1988, and Heck 1997. If they were, then the adequacy condition

express important facts about truth. (And, perhaps more to the point, important facts about the *truth conditions* of sentences.) But one's theory of truth needs to explain the T-sentences (and not the other way around, as deflationists maintain).

4.4. *Metaphysical neutrality*

One final core issue on which Tarski and primitivists are in complete agreement is the metaphysically neutral nature of truth. To accept Tarski's approach to truth is not to take on any metaphysical stance, such as realism or anti-realism: "we may accept the semantic conception of truth without giving up any epistemological attitude we may have had; we may remain naive realists, critical realists or idealists, empiricists or metaphysicians — whatever we were before. The semantic conception is completely neutral toward all these issues" (1944: 362). Likewise, taking truth to be a primitive, undefinable concept also requires no particular metaphysical stance.

By way of contrast, it's instructive to understand the traditional substantive theories of truth as building metaphysical stances directly into their accounts of truth.[22] For correspondence theories, something is true just in case it stands in the correspondence relation to some fact. As a result, truths are from the outset entangled with ontology. If *p* is to be true, there must be some entity E for it to correspond to. For correspondence theory, there is no truth without some accompanying ontology. This commitment might seem innocuous, but consider some problematic cases. Take the sentence 'There are no hobbits'. One might think that this sentence can be true without the benefit of anything existing: after all, it appears that it's true because *nothing* of a certain sort exists. But correspondence theorists must grant that there is some entity for it to correspond to, lest they allow some truths to be true in spite of not corresponding to anything, thereby giving up

---

would again be vacuous. See also Davidson 1990 and Patterson 2008b for further criticism of this reading.

22. Deflationary theories of truth, however, share the metaphysical neutrality of the semantic conception and primitivism.

their theory. Consider also mathematical and analytic truths: to what are they supposed to correspond? My intention is not to show that such questions cannot be answered; rather, it's that correspondence theories immediately face a distinctively metaphysical challenge that Tarski and primitivists can avoid. Any fully fleshed out correspondence theory must give a comprehensive account of the nature of the corresponding objects, and in so doing will impose a particular metaphysical view onto the nature of truth.

Similarly, coherence theorists and pragmatic theorists can be understood as incorporating partisan metaphysical views into their accounts of truth. If the truth of 'Snow is white' consists in a certain kind of coherence between our beliefs, then this highly suggests an (at least in part) idealist metaphysics that connects the whiteness of snow with features of what other beliefs we hold. If the truth of my beliefs is in part a function of what other beliefs I hold, then it's not the case that the truth of 'Snow is white' is a matter fully independent of me and my particular beliefs. Likewise, if the truth of 'Snow is white' is in part a function of what it's useful for me to believe, then the truth of the matter is again mind-dependent. Hence, coherence and pragmatic accounts of truth, just like correspondence theories, cannot embrace the metaphysical neutrality inherent to primitivism and Tarski's semantic conception of truth.

To say that Tarski's theory of truth is neutral is to allow that it *could* be paired with the metaphysical ideas underlying the traditional substantive theories of truth, if one were so inclined. Hence, Tarski's work on truth does not show that the metaphysical concerns of interest to the traditional theorists of truth are nonsense, or that they disappear once we adopt the semantic conception. It's just that these concerns are better explored not from the perspective of the theory of *truth*, but from the theory of *truthmaking* (or some other metaphysical approach). According to Tarski, any adequate theory of truth will usher forth the equivalences like

'Socrates is a philosopher' is true if and only if Socrates is a philosopher.

What we have here is one sentence (from the language doing the analyzing) being used to give the necessary and sufficient conditions for the truth of another (not necessarily distinct) mentioned sentence (from the language being analyzed). But to say that 'Socrates is a philosopher' is true if and only if Socrates is a philosopher is to say nothing whatsoever about what it is, if anything, that makes 'Socrates is a philosopher' true. What we can't say is this: Socrates is a philosopher makes true 'Socrates is a philosopher'. That's ungrammatical nonsense. And we must be extremely cautious in saying that what makes 'Socrates is a philosopher' true is that Socrates is a philosopher. For what does 'that Socrates is a philosopher' refer to? It can't refer to a proposition, because propositions don't make true the sentences that express them. If it refers to something like a "fact" or "obtaining state of affairs", then we have entities that can serve as truthmakers; but notice that this view is now burdened with some serious metaphysical commitments, and to say that all sentences are made true by such entities takes one perilously close to a traditional correspondence theory.

Ultimately, the question of what, if anything, makes 'Socrates is a philosopher' true is a separate metaphysical question left unanswered by everything that Tarski says. Perhaps it's made true by a fact, or state of affairs (Armstrong 1997), or by a trope (Lowe 2007), or by Socrates under a counterpart relation (Lewis 2003). Perhaps it doesn't have a truthmaker at all (Lewis 2001). Perhaps it's made true by the relations that obtain between my beliefs, or by the utility to me that would exist were I to believe what the sentence says. Regardless of which metaphysical view is correct, we can all agree that the Tarskian equivalence above is true. Metaphysical disputes regarding what it is that makes truth bearers true are best explored from within the metaphysical enterprise of truthmaking, and not the theory of truth itself. [23] This

23. It's worth noting that Tarski's claim to metaphysical neutrality is in some tension with his claim to be doing justice to our classical conception of truth,

neutral perspective is yet another point of common ground between Tarski and primitivism.

5. Dissolving Objections

As we have seen, Tarski's semantic conception of truth has sometimes been thought to motivate correspondence theories, and sometimes deflationary theories. My alternative suggestion is that Tarski's work is best paired with the primitivist perspective on truth. Primitivists stress the indefinability of our most general concept of truth, its explanatory value, and its metaphysical neutrality. All of these features, together with primitivism's perspective on the T-sentences, are shared with Tarski's semantic approach to truth. Hence, although Tarski himself never says anything that commits him to primitivism, he says plenty of things that distinguish him from both the traditional substantive theories and deflationary theories. If Tarski were to accept any contemporary account of our general concept of truth, primitivism would be the view most acceptable to him.[24] Likewise, primitivists can find support for their views inside Tarski's work. Tarski argues that his own method of definition cannot be applied to the general notion of truth that most interests primitivists. Thus, Tarski leaves open the possibility that primitivism offers the best understanding of our concept of truth.

---

what I have called the "correspondence intuition". If the correspondence intuition has any metaphysical pull to it — and I believe that it does — then it ought not receive any support from a metaphysically neutral view. On my view, consistency with the T-sentences is necessary but not sufficient for capturing the correspondence intuition. For this reason, I argue in my 2011 that deflationists cannot capture the correspondence intuition by their conformity to the truth schema alone, *contra* Horwich 1990. What's needed, in addition, is a particular view about what *makes* truths true.

24. Patterson observes that "Tarski had no specific concerns at all about the content of the concept of truth itself" (2012: 139). On Patterson's view, Tarski's fundamental interest is not in analyzing our intuitive notion of truth, but in showing how to consistently and adequately introduce a truth predicate (that conforms to the intuitive notion) into certain kinds of languages. If so, then Tarski would have been officially neutral regarding primitivism; my contention is that the view is open and amenable to him, if he were to throw his hat into this particular philosophical ring.

Tarski's reticence regarding what we should say about our natural language understanding of truth has led several philosophers to protest that by restricting his positive views to formal languages, Tarski's work fails to be of any philosophical interest. Armed with our primitivist perspective on Tarski's views, we can confront these objections directly. For example, according to Max Black, "The philosophical relevance of [Tarski's] work will depend upon the extent to which something similar [to his formal language definitions] can be done for colloquial English" (1948: 56). But this response begs the question against primitivism: Tarski's work on the limits of definability encourages the thought that truth as conceived in ordinary thought and speech is not definable, and this fact is of profound philosophical relevance. The question of which of our concepts are foundational and fundamental is an important one; Black, to the contrary, assumes that showing something to be indefinable and primitive cannot be philosophically relevant.

Other philosophers have also objected to Tarski's theory of truth on the grounds that it has little to offer, philosophically speaking. In the most infamous case, Putnam declares that "As a philosophical account of truth, Tarski's theory fails as badly as it is possible for an account to fail" (1985: 64), and denies that what Tarski defines "is in any way similar in meaning" to our intuitive use of truth (1983: 37). Putnam's worry derives from his view of Tarski's truth predicates, which (according to Putnam) apply to sentences regardless of what they mean. In effect, whether or not a sentence possesses Tarskian "truth" doesn't depend on what it means, and so Tarskian "truth" can't be real truth, which does depend on meaning.

This is not the place to evaluate Putnam's objection in full;[25] what we may appreciate instead is how a Putnam-style worry simply dissolves when we view Tarski's work through the primitivist lens. Tarski's ambitions did not include offering anything like a reductive definition in the mold of the traditional theories of truth. One should not turn to

Tarski in order to learn the ultimate nature of truth, for Tarski may well agree with primitivists that no such theory can be provided.[26] However, this doesn't show that Tarski's work fails as a "philosophical" theory, for that response presupposes that the only philosophical theory of truth worth having is a reductive, non-primitivist one. We may grant to Putnam that equivalences like "'Snow is white' is true-in-L if and only if snow is white' (even if read in Putnam's problematic way, as necessary logical truths) do not reveal at all the real nature of truth. But that was never their purpose. Tarski offers his definitions of truth in order to provide semanticists, logicians, mathematicians, interpreters of science, and others predicates that are safe from paradox and metaphysical baggage, and yet remain true to our basic understanding of the notion of truth. The T-sentences, again, provide a kind of "check" that his definitions conform to that basic understanding; they are not themselves supposed to be defining our basic concept of truth. Putnam's objection misunderstands the role that T-sentences play in the theory of truth, and supposes Tarski to be up to something that he doesn't think can be done.

Etchemendy (1988) has also voiced a number of influential criticisms of Tarski's work. He argues that Tarski's project of defining truth is actually at odds with the project of understanding semantics, despite appearances to the contrary. Etchemendy argues that Tarski, in order to define a notion of truth that does not succumb to paradox, effectively stipulates a definition of truth, and any such stipulative definition cannot contain the empirical information essential to the semantic theory of a language. Tarski's definitions will usher forth equivalences such as "'Snow is white' is true (in the Tarskian sense) if and only if snow is white', guaranteeing that truth (in the Tarskian sense) gets the right extension. But this equivalence doesn't give us any semantic information about the English sentence 'Snow is white'. To do that, we need not the (supposedly) logical truth

---

25. For further discussion, see Davidson 1990, Raatikainen 2003 and 2008, and Patterson 2008a and 2008b. See also the closely related discussions in Soames 1984, Etchemendy 1988, and Heck 1997.

26. Putnam briefly suggests that the primitivist line is not open to Tarski (1985: 72), but he does not say why. Perhaps he has in mind the sorts of considerations Etchemendy considers, as discussed below.

above involving (Tarskian) truth, but rather the equivalence ''Snow is white' is true (in the ordinary sense) if and only if snow is white'. Unlike Tarski's equivalence (according to Etchemendy), this second equivalence expresses contingent, empirical information about the sentence 'Snow is white'. Tarski's equivalence might help define a notion that dodges paradox, but it won't tell us anything about the semantics of natural language.

Now, Etchemendy goes on to claim that one can move from Tarski's "list-like" definition of truth for a language to a semantic theory for that language by employing a primitive notion of truth (1988: 59). On this approach, we take the additional theoretical step of holding that any particular truth predicate defined in Tarskian fashion expresses our ordinary notion of truth, and thus hold that anything true (in the Tarskian sense) is true (in the ordinary sense). So it might appear that one can use (a properly supplemented form of) Tarski's project to aid in semantics. But the appearances here are deceptive, says Etchemendy, and "little more than a fortuitous accident" (1988: 52). For Etchemendy argues that amending Tarski in this primitivist fashion is in conflict with Tarski's ambitions. Tarski wants an eliminative notion of truth — and so one that can be properly inoculated from paradox — and the attempt to make Tarski's theory philosophically informative ends up relying on an uneliminated primitive conception of truth.

Even setting aside my aforementioned doubts that Tarski reads the T-sentences as being purely stipulative logical truths, I do not believe that we need to follow Etchemendy in seeing the primitivist maneuver as being antithetical to Tarski's ambitions. As we have seen, Tarski's interest was never in *replacing* our ordinary conception of truth with the kind of definitions he offers. Rather, Tarski's definitions work in *conjunction* with our ordinary conception of truth. We know that Tarski's definitions are successful only if they are materially adequate, in which case they entail all the T-sentences. For the T-sentences to provide an *independent* check on Tarski's definitions, they must be expressing important facts about our ordinary conception of truth (or, at

least, about the truth conditions of our sentences), and not vacuous logical truths. So whereas Etchemendy sees a prior, primitivist conception of truth as somehow conflicting with Tarski's project, the correct interpretation is that Tarski's project relies on there being such a prior conception. As I have argued, primitivism about that prior conception is the view that is most consonant with Tarski's project.

What remains to be seen is whether primitivists can convince Tarski that our use of truth in ordinary thought can be consistent. Tarski worried that our use of truth in natural language was inconsistent, though he stopped short of definitively drawing that conclusion. So the question arises how the primitivist may approach the semantical paradoxes as they arise in natural language. That topic — primitivism and paradox — is one that we shall have to take up elsewhere.[27, 28]

References

Aristotle. 1966. *Metaphysics*. Trans. Hippocrates G. Apostle. Bloomington: Indiana University Press.

Armstrong, D. M. 1997. *A World of States of Affairs*. Cambridge: Cambridge University Press.

Asay, Jamin. 2011. *Truthmaking, Truth, and Realism: New Work for a Theory of Truthmakers*. Doctoral dissertation. Chapel Hill: University of North Carolina.

——. Forthcoming. Against *truth*. *Erkenntnis*.

Bar-On, Dorit, and Keith Simmons. 2007. The use of force against deflationism: assertion and truth. In *Truth and Speech Acts: Studies in*

27. See also Patterson 2006b for thoughts on how Tarski may approach the semantic paradoxes for natural language.

28. My thanks go to the referees for *Philosophers' Imprint* for their thorough, helpful comments. I'd also like to thank Keith Simmons, both for his careful reading of the paper, and also for his patience in helping me develop my thoughts on Tarski. Any interpretive errors, of course, are entirely mine. Material from this paper will be incorporated into my monograph *The Primitivist Theory of Truth*, forthcoming with Cambridge University Press.

the Philosophy of Language, eds. Dirk Greimann and Geo Siegwart, 61–89. London: Routledge.

Black, Max. 1948. The semantic definition of truth. *Analysis* 8: 49–63.

David, Marian. 1994. *Correspondence and Disquotation: An Essay on the Nature of Truth*. Oxford: Oxford University Press.

Davidson, Donald. 1969. True to the facts. *Journal of Philosophy* 66: 748–764.

——. 1973. In defense of convention T. In *Truth, Syntax and Modality: Proceedings of the Temple University Conference on Alternative Semantics*, ed. Hugues Leblanc, 76–86. Amsterdam: North-Holland Publishing Company.

——. 1990. The structure and content of truth. *Journal of Philosophy* 87: 279–328.

——. 1996. The folly of trying to define truth. *Journal of Philosophy* 93: 263–278.

——. 1999. The centrality of truth. In *Truth and Its Nature (If Any)*, ed. Jaroslav Peregrin, 105–115. Dordrecht: Kluwer Academic Publishers.

DeVidi, David, and Graham Solomon. 1999. Tarski on "essentially richer" metalanguages. *Journal of Philosophical Logic* 28: 1–28.

Etchemendy, John. 1988. Tarski on truth and logical consequence. *Journal of Symbolic Logic* 53: 51–79.

Fernández Moreno, Luis. 2001. Tarskian truth and the correspondence theory. *Synthese* 126: 123–147.

Field, Hartry. 1986. The deflationary conception of truth. In *Fact, Science and Morality: Essays on A. J. Ayer's* Language, Truth and Logic, eds. Graham Macdonald and Crispin Wright, 55–117. Oxford: Basil Blackwell.

——. 1994. Deflationist views of meaning and content. *Mind* (New Series) 103: 249–285.

——. 2008. *Saving Truth from Paradox*. Oxford: Oxford University Press.

Frege, Gottlob. 1956. The thought: a logical inquiry. Trans. A. M. Quinton and Marcelle Quinton. *Mind* (New Series) 65: 289–311.

García-Carpintero, Manuel. 1999. The explanatory value of truth theories. In *Truth and Its Nature (If Any)*, ed. Jaroslav Peregrin, 129–148. Dordrecht: Kluwer Academic Publishers.

Haack, Susan. 1976. 'Is it true what they say about Tarski?' *Philosophy* 51: 323–336.

——. 1978. *Philosophy of Logics*. Cambridge: Cambridge University Press.

Halbach, Volker. 2011. *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Heck, Richard G., Jr. 1997. Tarski, truth, and semantics. *Philosophical Review* 106: 533–554.

Horsten, Leon. 2011. *The Tarskian Turn: Deflationism and Axiomatic Truth*. Cambridge: MIT Press.

Horwich, Paul. 1982. Three forms of realism. *Synthese* 51: 181–201.

——. 1990. *Truth*. Oxford: Basil Blackwell.

——. 2005. A minimalist critique of Tarski on truth. In *Deflationism and Paradox*, eds. Jc Beall and Bradley Armour-Garb, 75–84. Oxford: Clarendon Press.

Jennings, Richard C. 1987. Is it true what Haack says about Tarski? *Philosophy* 62: 237–243.

Ketland, Jeffrey. 1999. Deflationism and Tarski's paradise. *Mind* (New Series) 108: 69–94.

Keuth, Herbert. 1978. Tarski's definition of truth and the correspondence theory. *Philosophy of Science* 45: 420–430.

Kirkham, Richard L. 1992. *Theories of Truth: A Critical Introduction*. Cambridge: MIT Press.

Künne, Wolfgang. 2003. *Conceptions of Truth*. Oxford: Clarendon Press.

Leeds, Stephen. 1978. Theories of reference and truth. *Erkenntnis* 13: 111–129.

Lewis, David. 1983. New work for a theory of universals. *Australasian Journal of Philosophy* 61: 343–377.

——. 2001. Truthmaking and difference-making. *Noûs* 35: 602–615.

——. 2003. Things qua truthmakers. In *Real Metaphysics: Essays in Honour of D. H. Mellor*, eds. Hallvard Lillehammer and Gonzalo Rodriguez-Pereyra, 25–42. London: Routledge.

Lowe, E. J. 2007. Truthmaking as essential dependence. In *Metaphysics and Truthmakers*, ed. Jean-Maurice Monnoyer, 237–259. Frankfurt: Ontos Verlag.

Martin, Robert L. 1968. On Grelling's paradox. *Philosophical Review* 77: 321–331.

Moore, George Edward. 1899. The nature of judgment. *Mind* (New Series) 8: 176–193.

Niiniluoto, Ilkka. 1999. Tarskian truth as correspondence — replies to some objections. In *Truth and Its Nature (If Any)*, ed. Jaroslav Peregrin, 91–104. Dordrecht: Kluwer Academic Publishers.

Patterson, Douglas Eden. 2006a. Tarski on the necessity reading of Convention T. *Synthese* 151: 1–32.

——. 2006b. Tarski, the liar, and inconsistent languages. *Monist* 89: 150–177.

——. 2008a. Tarski's conception of meaning. In his *New Essays on Tarski and Philosophy*, 157–191. Oxford: Oxford University Press.

——. 2008b. Truth-definitions and definitional truth. *Midwest Studies in Philosophy* 32: 313–328.

——. 2010. Truth as conceptually primitive. In *New Waves in Truth*, eds. Cory D. Wright and Nikolaj J. L. L. Pedersen, 13–29. New York: Palgrave Macmillan.

——. 2012. *Alfred Tarski: Philosophy of Language and Logic*. New York: Palgrave Macmillan.

Popper, Karl R. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge.* New York: Harper and Row.

——. 1979. Is it true what she says about Tarski? *Philosophy* 54: 98.

Putnam, Hilary. 1983. On truth. In *How Many Questions? Essays in Honor of Sidney Morgenbesser*, ed. Leigh S. Cauman, Isaac Levi, Charles Parsons, and Robert Schwartz, 35–56. Indianapolis: Hackett Publishing Company.

——. 1985. A comparison of something with something else. *New Literary History* 17: 61–79.

Raatikainen, Panu. 2003. More on Putnam and Tarski. *Synthese* 135: 37–47.

——. 2008. Truth, meaning, and translation. In *New Essays on Tarski and Philosophy*, ed. Douglas Patterson, 247–262. Oxford: Oxford University Press.

Ray, Greg. 2003. Tarski and the metalinguistic liar. *Philosophical Studies* 115: 55–80.

——. 2005. On the matter of essential richness. *Journal of Philosophical Logic* 34: 433–457.

Resnik, Michael D. 1990. Immanent truth. *Mind* (New Series) 99: 405–424.

Russell, Bertrand. 1904. Meinong's theory of complexes and assumptions (III.). *Mind* (New Series) 13: 509–524.

Schantz, Richard. 1998. Was Tarski a deflationist? *Logic and Logical Philosophy* 6: 157–172.

Simmons, Keith. 1990. The diagonal argument and the Liar. *Journal of Philosophical Logic* 19: 277–303.

——. 1993. *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. Cambridge: Cambridge University Press.

——. 2009. Tarski's logic. In *Handbook of the History of Logic, Volume 5: Logic from Russell to Church*, eds. Dov M. Gabbay and John Woods, 511–616. Amsterdam: North-Holland.

Soames, Scott. 1984. What is a theory of truth? *Journal of Philosophy* 81: 411–429.

——. 1999. *Understanding Truth*. New York: Oxford University Press.

Stoutland, Frederick. 1999. Do we need correspondence truth? In *Truth and Its Nature (If Any)*, ed. Jaroslav Peregrin, 81–90. Dordrecht: Kluwer Academic Publishers.

Strawson, P. F. 1992. *Analysis and Metaphysics: An Introduction to Philosophy*. Oxford: Oxford University Press.

Tarski, Alfred. 1944. The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research* 4: 341–376.

——. 1956a. The concept of truth in formalized language. In his *Logic, Semantics, Metamathematics: Papers From 1923 to 1938*, trans. J. H. Woodger, 152–278. Oxford: Clarendon Press.

——. 1956b. The establishment of scientific semantics. In his *Logic, Semantics, Metamathematics: Papers From 1923 to 1938*, trans. J. H. Woodger, 401–408. Oxford: Clarendon Press.

Textor, Mark. 2010. Frege on judging as acknowledging the truth. *Mind* (New Series) 119: 615–655.

Williams, Michael. 1999. Meaning and deflationary truth. *Journal of Philosophy* 96: 545–564.

Woleński, Jan. 1993. Two concepts of correspondence. *From the Logical Point of View* 2.3: 42–55.