

**TRUTH & TRANSCENDENCE:
Turning the Tables on the Liar Paradox**

Gila Sher

I. A Methodological Turnaround.

Confronting the Liar Paradox is commonly viewed as a prerequisite for developing a theory of truth. As soon as the truth theorist accepts one of the most minimal principles of truth, the *equivalence principle*, which in one of its forms is often formulated (schematically) by

(E) $\langle P \rangle$ is true iff (if and only if) P ,

where “ P ” stands for any sentence and “ $\langle P \rangle$ ” stands for a name of that sentence, the Liar Paradox arises. And it is only after the truth theorist incorporates some device (like a Tarskian hierarchy or a Kripkean grounding process) for blocking the paradox that he is entitled to continue developing his theory.

In this paper I would like to turn the tables on this traditional conception of the relation between the Liar Paradox and the theory of truth. I would like to show that the theorist of truth need not worry about the Liar Paradox in developing her theory, that if she focuses deeply enough on the “material” adequacy of her theory (i.e., whether it adequately accounts for the *nature* of truth) as distinct from its “formal” adequacy (i.e., whether it is a consistent theory), the Liar challenge is unlikely to arise for her theory at all.¹

In approaching the Liar challenge in this way, we are treating the theory of truth like most other theories. Consider a physical, psychological, geometrical, or moral theory. Normally,

¹ The contrast between “material” and “formal” is intended to invoke the similar contrast in Tarski (1933), but my notion of “material adequacy” is not limited to “extensional correctness”, as Tarski’s notion is sometimes thought to be. “Material”, in the sense intended here, has the connotation of “internal to the subject matter of the theory”, whereas “formal” has the connotation of “having to do with formal matters that are extraneous to the subject matter of the theory”. (When the subject matter of a theory is itself formal in the sense of dealing with, say, logical or set theoretical issues, the contrast between “material” and “formal” in the present sense is reduced or altogether disappears.)

the theorist in any of these fields is aware that his theory must satisfy certain norms of formal adequacy. But his focus in constructing his theory is not on this matter. His focus is on the *material adequacy* of his theory, on its success in giving a correct, comprehensive, and explanatory account of the physical structure of the world, human psychology, the geometry of space, the grounds of human morality, etc. Once his theory is completed (or reaches a temporary state of completion), the theorist has to verify, to the best of his ability, its formal consistency. Of course, if it turns out that his theory is inconsistent, this is a serious matter and the theorist must reexamine the theory: its material principles, its logical-linguistic framework, its background assumptions, etc. But under normal circumstances, the need to guard against inconsistency does not dominate, or even noticeably affect, the construction of our theories. Our theories are constructed based on material considerations, and the consistency check is just that, a check. In contrast, the theorist of truth, on the approach we are challenging, cannot go about his business – the construction of a correct, comprehensive, and informative theory of truth – in this way. Here, the issue of consistency is a major concern right from the beginning, and he is not entitled to proceed in developing the material content of his theory before taking care of its formal adequacy, or so the common wisdom says.

My claim in this paper is that the task of constructing a theory of truth is methodologically more similar to that of constructing a physical, psychological, geometrical, or moral theory than it is commonly thought to be. More specifically, if the theorist of truth builds a *materially good* theory (choosing some – any – reasonable logical-linguistic framework for working in), then the Liar challenge is unlikely to arise at all for her theory.

The key to this turnaround is methodological, and one way to explain it is by saying that according to the prevalent conception of the theory of truth, constructing a theory is like constructing a logical inference. But in fact *constructing a theory is very different from constructing a logical inference*. Let me explain. If you think of your theory of truth as a logical inference, where each of its principles is an independent premise and the question is whether

these premises entail a contradiction, then you conclude that if the equivalence principle, E, leads to a paradox, then the equivalence principle together with any other material principle of truth, M, will form a conjunction that also leads to that paradox:

$$(1) \quad [E \Rightarrow \text{Paradox}] \Rightarrow [E \ \& \ M \Rightarrow \text{Paradox}].^2$$

If this were correct, it would justify the requirement that we take care of the Liar Paradox as soon as we accept the equivalence principle. But this view of a theory is wrong. A theory – say, a philosophical theory – is not like a set of premises in a logical inference, and the process of building a philosophical theory is not that of creating a conjunction of independent principles. When you build a philosophical theory, each added principle constrains, and is constrained by, the other principles. That is, it affects, or can in principle affect, their scope and content, hence their consequences, and is similarly affected by them. As a result, if “ \boxtimes ” is the operation of placing two principles, P_1 and P_2 , together in a theory, then:

$$(2) \quad [P_1 \Rightarrow \text{Paradox}] \Rightarrow [P_1 \boxtimes P_2 \Rightarrow \text{Paradox}].$$

The difference between (2) and (1) is due to the difference between \boxtimes and $\&$. The process of constructing a theory is such that the addition of a new principle to the theory in principle involves the *updating* of all its principles, old and new. We may say that $P_1 \boxtimes P_2$ it is more like $[(P_1 \uparrow P_2) \ \& \ (P_2 \uparrow P_1)]$, where “ $X \uparrow Y$ ” stands for “X as updated by Y”, than like $P_1 \ \& \ P_2$. And the fact that for an arbitrary consequence C

$$(3) \quad [P_1 \Rightarrow C] \Rightarrow [P_1 \uparrow P_2 \Rightarrow C]$$

explains why

$$(4) \quad [P_1 \Rightarrow C] \Rightarrow [P_1 \boxtimes P_2 \Rightarrow C].$$

My suggestion is that there is some material principle of truth of, M, a principle arrived at by investigating the *nature of truth itself*, such that while

² Here and below the arrows represent logical, or metalogical, consequence relations, or more broadly, monotonic consequence relations.

- (5) $E \Rightarrow$ Liar Paradox,
 (6) $E \boxtimes M \Rightarrow$ Liar Paradox.

Which material principle of truth is M? There could, in principle, be multiple such principles. In this paper I will focus on one such principle, “IMMANENCE” (upper case), introduced in Sher (2004). More specifically, I will focus on two subprinciples of IMMANENCE, “immanence” (lower case) and “transcendence”, which, with the help of a third subprinciple, “normativity”, do the active work of preventing the Liar Paradox from arising. This pair of *material* principles says that truth is inherently *hierarchical*, and as such it is not susceptible to the Liar Paradox. We might say that the addition of IMMANENCE to Equivalence *updates* the latter in a way that blocks the Liar Paradox.

This turnaround in our understanding of the relation between the Liar Paradox and the theory of truth is not just significant by itself, but it has significant consequences for issues widely discussed in the literature:

1. It undermines a major criticism of Tarski’s theory of truth, namely, that the hierarchical structure it attributes to truth is ad hoc. More precisely:
2. It shows that the hierarchical element in Tarski’s and others’ theories of truth can be justified on material and not just on formal grounds, i.e., that it has deep roots in the material nature of truth, and is not just a technical device for dealing with a formal problem.
3. It further shows that many anti-Tarskian theories of truth share a hierarchical element with the theory they reject.
4. It disconnects both the Liar Paradox and the hierarchical conception of truth from the question of which logic is the right “logic of truth”, bivalent logic or non-bivalent logic. This it does by showing that the hierarchical element in truth is independent of bivalence and compatible with non-bivalence.
5. It also disconnects the Liar Paradox and the hierarchical conception of truth from the question of natural language vs. artificial language.
6. It draws new lessons from the existence of non-Tarskian solutions to the Liar Paradox.

7. It points to an additional advantage of a *substantivist* over a *deflationist* theory of truth besides the straightforward advantages of greater depth, informativeness, and explanatory power. Investigating the nature of truth beyond the minimalist equivalence principle, it shows, might save us from problems – including formal problems – that arise for deflationist or, more generally, “bare bones” theories of truth.

I should emphasize, however, that saying we do not need to consider the Liar Paradox prior to the construction of a material theory of truth does not minimize the value of facing the Liar Paradox on its own. On the contrary. Confronting the Liar Paradox on its own, i.e., independently of the development of a materially adequate theory of truth that might block it, has led to immense progress in our understanding of such central topics of philosophical inquiry as language, logic, semantics, definition, circularity, self-reference, vagueness, contextuality, revision, truth itself, and more.

II. A *Substantivist Theory of Truth*

Our main task is to present a material principle of truth that blocks the Liar Paradox. But this principle does not exist in a vacuum. It belongs to a certain theory, with its underlying methodology, goals, and perspective.³ Among the underlying features of this theory are (i) a substantivist orientation, (ii) a holistic, or more precisely, a “foundational-holistic” methodology, and (iii) a focus on the cognitive-epistemic role of truth. Let me briefly explain the nature of and motivation for these underpinnings:

(i) *Substantivist Orientation*. The motivation for pursuing a substantive theory of truth in contrast to a trivial, thin, or a bare bones deflationist theory of truth is straightforward and common sensical. If the subject matter of truth is thin and trivial, there is no need to develop a *theory* of this subject matter. If, on the other hand, the subject matter of truth is rich and philosophically salient, it requires a *substantive theory*. In this respect, too, I regard the field of

³ For various aspects of this theory, including further discussion of issues that I will only briefly gloss over below, see Sher (1999a, 2004, 2012, and Forthcoming 2,3).

truth as similar to all other fields of knowledge, both within and outside philosophy. If the physical structure of the world is a rich and important subject matter, it requires a substantive – deep, thorough, rich, informative, explanatory – theory. And the same holds for knowledge, logical reasoning, morality, and other philosophical subject matters, including truth.

But the development of substantive theories of many philosophical subject matters faces special challenges, both challenges due to the nature of these subject matters and challenges due to philosophers' conception of the structure of philosophical theories of such subject matters. Subject matters like truth and knowledge are characterized by extraordinary breadth, diversity, and complexity; yet one influential conception of a *philosophical theory* of these subject matters is that of a single and simple *definition* or *necessary and sufficient condition*. In the case of knowledge such a “theory” used to take the form of a definition like “Knowledge =_{Df} True, justified belief” or a necessary and sufficient condition like “x knows that P iff x has a true, justified belief that P”. In the case of truth, it often takes the form of the *equivalence* biconditional, under one formulation or another. The combination of these two circumstances – the breadth of many philosophical subject matters and philosophers' narrow conception of a theory of these subject matters – introduces a nontrivial challenge to substantivist philosophers.

In the case of truth, this challenge may appear insurmountable. Truth is a rich and philosophically significant subject matter, interwoven in many areas of our life, from the epistemic and theoretical to the moral and practical, and this creates a serious methodological problem for at least one central branch of the theory of truth, that dealing with *truth conditions*. The truth conditions of sentences (thoughts, beliefs, propositions, theories, bodies of knowledge, etc.)⁴ are tied up with their content, and the enormous diversity of *types of content* creates a plurality of *types of truth conditions*. (For example, the truth conditions of a discourse about causally accessible everyday physical objects are likely to be quite different in kind from those of a highly abstract discourse – discourse about causally inaccessible physical phenomena,

⁴ I intentionally construe potential truth-bearers very broadly.

mathematical laws, moral principles, etc.) A major challenge for the theorist of truth is dealing with this diversity. The traditional conception of the theory of truth as comprised of a single and simple definition or necessary and sufficient condition, one that captures the one and only common denominator of all truths, or the one and only essence of truth, or the one and only necessary and sufficient condition for any sentence to be true, makes this challenge impossible to meet. If truth is too diverse, too complex, and too multidimensional to be exhausted by a single common denominator, then a *substantive* theory of truth is unfeasible.

But this conclusion, I have shown in Sher (1999a, 2004), is unwarranted. If the conception of a theory of truth as a single and simple definition is untenable, it can and should be replaced by that of a *cluster* of interconnected *principles* of truth (of various degrees of complexity and generality), a conception widely accepted by theorists in most fields of knowledge. Within philosophy, such a conversion has been successfully accomplished in some fields, notably epistemology, where during a certain period in the 20th century many theorists viewed their task as producing a definition of, or a necessary and sufficient condition for, “x knows that P”, and today only few follow in their footsteps.

Indeed, the diversity challenge is not unique to philosophy. A similar challenge arises for natural science as a whole, that is, as our *theory of nature*. The diversity, complexity, and multidimensionality of nature pose a serious challenge to the theorist of nature. This challenge has been widely discussed by philosophers and scientists under the rubric of “the disunity of science” and, not surprisingly, many of its lessons are independent of the empirical nature of natural science. In Sher (2004) I suggested that philosophers ought to adopt the same common sense guidelines for dealing with this challenge as those devised by scientists. The gist of this idea is concisely captured by Dyson when he says that “every theory needs for its healthy growth a creative balance between unifiers and disunifiers” (1988:47).

In the case of truth, my suggestion is that we think of a substantive theory as a body of substantive principles – some more general, others more particular, some simpler, others more

complex, some manifesting greater inner unity, others greater diversity, but all interconnected. This strategy frees us to search for, rather than legislate in advance (e.g., in the form of a definition by fiat), the material principles of truth, including principles that involve a significant adjustment, or updating, of other principles.

(ii) *Foundational-Holistic Methodology*. Another challenge facing the theorist of truth, along with theorists of other philosophical subject matters, is the “foundational” challenge. Philosophical theories often deal with very basic subject matters that call for a *foundational* treatment. But foundational studies in philosophy have come upon serious difficulties and many philosophers view them as doomed to failure. In Sher (2010 and Forthcoming 1) I have suggested that this is due to the fact that the *foundational* project is commonly associated with a self-defeating traditional methodology, the so called *foundationalist* methodology. The foundationalist methodology sets strict, unsatisfiable requirements on the foundational project, including a strict ordering of all areas of knowledge. To conduct a foundational study of an area of knowledge X we must limit ourselves to cognitive resources generated in areas lower than X. This means that it is impossible to generate tools for a foundational study of the most basic areas of knowledge, those lying at the base of the ordering. The foundational study of such areas is bound to involve either circularity or resources generated higher-up in the ordering, and as such it violates a fundamental principle of the foundationalist methodology. This has led many philosophers to shift from the foundationalist to the *holistic* methodology, but holism is commonly believed to involve a complete renouncement of the *foundational* (and not just the *foundationalist*) project or else acquiescence to a very limited (e.g., a coherentist or a narrowly naturalistic) foundational project.

In the above mentioned works I argue that holism is perfectly compatible with a robust, noncoherentist, non-narrowly-naturalistic foundational project. The foundational project, rightly conceived, is the project of substantively and critically studying the main principles of some basic subject matter, with special emphasis on *explanation* and *justification*. Now, if we think of this

project as one that is pursued in ways appropriate for humans (rather than, say, for gods), then we are seeking not absolute certainty or instantaneous knowledge (achieved by a flash of intuition), but progress achieved by probing inquiry, carefully monitored imagination and insight, smart decisions, experimentation, openness to criticisms, willingness to institute revisions, and so on. Thus understood, the foundational project does not require a strict or a rigid methodology. Rather, it favors a flexible and dynamic methodology, one that sanctions less-than-perfect and temporary tools, back and forth movement, shifts in perspectives, and so on – i.e., the kind of methodology that is manifested in the holistic metaphor of Neurath’s boat (minus coherentism). On this interpretation, Neurath’s boat is a boat on a mission, a mission to study the sea and its residents, and to achieve this goal its occupants are ready to use any tools available to them flexibly and constructively yet also thoughtfully and critically. It is this methodology, which I call “foundational holism”, that makes a substantive, foundational study of basic philosophical subject matters like truth possible.

(iii) *Cognitive-Epistemic Focus*. Every philosophical theory approaches its subject matter from some perspective. The theory of truth considered in the present paper approaches truth from a *cognitive-epistemic* perspective.⁵ This does not mean that it identifies truth conditions with epistemic conditions (e.g., justification conditions). On the contrary; it regards the task of assigning distinctly *truth* conditions to sentences and theories as *integral* to the epistemic project. Truth is fundamental to the human cognitive-epistemic project due to a combination of circumstances characterizing the human condition: strong cognitive interests (both practical and intellectual), complex world, and a mixed assortment of cognitive resources (limited in some respects, rich and intricate in others). This complex situation creates an abiding need for *reality checks* by human cognizers, and that, in turn, requires a standard of truth for human thoughts

⁵ “Cognitive” and “epistemic”, or “cognition” and “knowledge”, are used here as cognate notions, with “cognition” being the broader, weaker, and less specific of the two. While “knowledge” connotes *success* in cognition, “cognition” connotes mainly an *attempt* at acquiring knowledge (information).

(beliefs, theories). In other words: It is due to the everpresent threat of a *gap* between our *thought* on the one hand and *reality* on the other that the question of truth is so crucial for us.⁶ Had life been so hospitable that we needed no information about the world, had we no intellectual interest in the world (no desire to have a theoretical understanding of the world), were we incapable of cognitively diverging from or going beyond what is actually the case, had we no imagination, no drives, interests, or motives that blocked or distorted information about the world, then a standard, a concept, and a theory of truth would have been of little cognitive use for us. But since none of this is the case, the question of truth, the question whether our thoughts⁷ *measure up to reality*, always arises for us and is of great importance to us.

III. *The IMMANENCE Thesis and Its Three Constituting Principles: Immanence, Transcendence, and Normativity.*

The IMMANENCE thesis addresses a semi-Kantian question about truth, namely: Under what cognitive conditions does truth emerge as a fundamental standard of correctness for human thought? Its answer is that *for truth to arise, three conditions on human cognition have to be satisfied: immanence, transcendence, and normativity*. Put differently: *Truth emerges in the intersection of three basic modes of human cognition: immanence, transcendence, and normativity*.

Immanence. The first condition for the emergence of truth as a fundamental standard for human thought is the ability, and practice, of *directing our gaze at the world, or at some things in the world, and saying something about them, or attributing some property, relation, or state to them*. The mode of thought we use in satisfying this condition I call the “immanent mode” and thoughts exhibiting the characteristics of this mode – *immanent thoughts*. Immanence, thus, is

⁶ I use “reality” and “world” as synonyms.

⁷ Here and in the remainder of this paper I use “thought” as a general term that covers sentence, statement, belief, judgment, cognition, theory, body of knowledge, thought proper, etc.

both a *mode* and a *property* of thought: a mode of thought as an act, and a property of thought as an object.

My use of “immanence” for this mode/property is influenced by Quine. In some of his writings (e.g., 1981: 21-2) Quine says that to speak immanently is to speak from *within a theory*. But speaking from within a theory is, typically, saying something about things *outside the theory*, things *in the world*. In my own use, speaking immanently is speaking *in the way* one typically speaks when one speaks from within a theory, namely, speaking about some subject matter, attributing properties/relations to some objects, or saying how the world is. Immanence, thus, exhibits a basic dialectic of human theories: theories are human creations, yet they are focused on something external to them – the world, in a broad sense of the world. Accordingly, “immanent” connotes “being *internal* to a *theory*”, but “being internal to a theory” connotes “being directed at something *external* to the theory”.⁸

Our conception of immanence is also related to ideas by other philosophers: it is related to some philosophers’ idea of *intentional* thought (see, e.g., Siewert 2006), to Frege’s view that “in every judgment ... a step is made from the level of propositions to the level of the *nominata* (the objective facts)” (1892: 91)⁹, to James’ statement that “human thought appears to deal with objects independent of itself” (1890: 271), to Wittgenstein’s claim that “[t]he general form of propositions is: This is how things are” (1921: 4.5), etc.

My use of “immanence”, however, is also different from many uses of this term in the philosophical literature, including some aspects of Quine’s use of this term. For example, Quine (1970/86, 1986, 1995) restricts immanent statements to statements belonging to our mother

⁸ We could replace “immanence” with the Quinean expression “world oriented”, but to emphasize both the contrast, or complementarity, to transcendence and the inherent dialectic of *theories* we will stay with “immanence”.

⁹ This is especially clearly expressed in the Feigl translation. In the original: “[I]n jedem Urteil ... der Schritt von der Stufe der Gedanken zur Stufe der Bedeutungen (des Objectiven) geschehen ist”.

tongue, to a given object language, to scientific discourse, or to naturalistic discourse. My own conception of immanence does not impose any of these restrictions. Immanent thought, on my conception, is commonly translinguistic. The principles of general relativity, for example, are immanent in my sense, yet they do not belong to a specific language. Similarly, Kant's conditions for the possibility of knowledge are immanent, yet they are not part of a scientific or a naturalistic discourse.

The category of immanent thought determines the domain of *truth-bearers*, i.e., the range of thoughts for which truth serves as a standard and to which truth-properties (truth, falsehood, truth-indeterminacy, etc.) apply. Immanence sets no limit on the complexity of thoughts, on whether they address their subject matter literally or nonliterally, directly or indirectly. Given our broad conception of world, thoughts themselves are part of the world, and therefore the category of immanent thoughts includes thoughts directed at thoughts. The category of immanent thought is very broad, but not all thoughts are immanent. In particular, thoughts that do not intend, or do not succeed, in saying something about the world or in attributing a property (relation) to some things in the world are not immanent. To be immanent is to genuinely attribute a property to something in the world, something of the kind that the property in question can be attributed to, and this is not a trivial thing. Normally, the category of immanence encompasses all statements and theories of all genuine fields of knowledge as well as large parts of everyday discourse. But depending on the type of thought involved (the type of objects and properties involved) immanence might pose certain specific requirements, as we will see in the case of *truth-thoughts* – thoughts that attribute truth-properties to thoughts – below. Desirably, the boundaries of immanence are delineated in a systematic manner. But even before an adequate systematization is available, we can use our judgment to decide, in many particular cases, whether a given thought is

immanent.¹⁰ A nontrivial pre-systematic candidate for a non-immanent thought is “The number 2 is laughing”.¹¹

While immanence is a basic condition for truth, immanence by itself does not suffice to yield truth. To focus on the world, to say something about it, is not yet to approach it through the prism of *truth*.

Transcendence. A second condition for the emergence of truth as a fundamental standard for human cognition is *transcendence*. By “transcendence” I mean the ability, and actual practice, of *moving outside a given thought in order to reflect upon it, examine it, say something about it, ask and answer questions about it, set norms or standards for it, challenge it, attribute properties to it, and so on*. The transcendence required for truth is transcendence of *immanent* thoughts. Henceforth, we will understand by “transcendence” transcendence of this kind.

Transcendence has fallen into disrepute lately. To say that truth is transcendent, it is claimed, is tantamount to saying that we have a “God’s eye view” on language and the world. But “transcendence” does not need to have this connotation, and in our own use of this word it does not. Transcendence, as we understand it here, is not something mysterious or superhuman; rather, it is something quite simple and commonplace. Transcending an immanent thought, or a region of immanent thoughts, is casting a reflective look at it from a standpoint that holds it in view yet is located *within* the purview of human thought. In accordance with our foundational-holistic methodology, our conception of transcendence is holistic. Transcending a thought is finding a standpoint anywhere on Neurath’s boat from which we can see it. Transcendence of this kind is also dynamic. We transcend sociology in order to view it from a philosophical standpoint, and we transcend philosophy in order to view it from a sociological standpoint. What counts as *appropriate* transcendence varies according to task and circumstance. If our task is to describe the

¹⁰ Indeed, working outside a full systematization is sometimes advantageous, since the act of systematizing a theory may force us to introduce some ad hoc adjustments.

¹¹ But see fn. 16 below.

syntax of a given theory, then, as Gödel (1931) has shown, if the theory is sufficiently rich, we can view its syntax from within it. I.e., a theory itself can provide an adequate standpoint for examining some of its features. But if our task is different, we may need a standpoint *outside* it. While the principle of transcendence, like the principle of immanence, calls for a systematization, here too we are able to judge in many particular cases whether an appropriate transcendence is achieved prior to a full systematization.

Transcending an immanent thought or a domain of immanent thoughts is merely a *human* act, yet it is a cognitively *powerful* act. Gödel's completeness and incompleteness theorems, Church's thesis, Turing's proof of the unsolvability of the halting problem, the Löwenheim-Skolem theorem, Lindström's theorems, all testify to power of the transcendent mode in one of its forms – ascent to a metalanguage.

Transcendence is central to numerous fields of knowledge: psychology, sociology, many areas of philosophy, metalogic, metamathematics, etc. On the one hand, the very idea of metalogic or of philosophy (sociology, psychology) of, say, knowledge, requires transcendence; on the other, transcendence enables us to develop tools that are needed in these disciplines.

It is important to recognize that transcendence *does not conflict with immanence*. On the contrary, most transcendent thoughts are immanent. Their target is something in the world – human thoughts; they attribute properties to their target thoughts, relate them to other thoughts as well as to things other than thoughts, and so on. As such, they are *genuinely* immanent. There are many types of transcendent thoughts, and one of these is the *truth-thought* – a thought that attributes a truth-property to some thought(s). To be an *admissible* truth-thought, i.e., a truth-thought that is admissible as a truth-bearer, a truth-thought must be both *immanent* and *transcendent*, and its object(s) – the thought(s) to which it attributes a truth property – must also be *immanent*. We will call this interplay between immanence and transcendence the “immanence–transcendence complementarity”.

Immanence and transcendence by themselves, however, are still not sufficient for truth. Although they can accommodate truth-thoughts, they are independent of such thoughts. By transcending an immanent discourse either to a higher level of discourse or to a standpoint on Neurath's boat that has it in view, we can do many things that are not related to truth: we can ask questions of a variety of kinds about these thoughts (e.g., are they mathematical thoughts), attribute properties and relations of multiple types to them (e.g., being thoughts about something funny), set standards (norms) of different sorts for them (e.g., clarity), doubt, challenge, justify, refute them on diverse grounds, enjoy them or be disgusted by them, etc. Truth arises when we ask questions of a special kind about immanent thoughts and set standards (norms) of a special kind for such thoughts. This requires a third mode of thought.

Normativity. A third condition for the emergence of truth as a fundamental standard for human cognition is the ability to *engage in normative activities* and the *actual engagement in such activities*. The mode of thought characteristic of such engagement I call the "normative mode of thought". The normative mode of thought is a mode of questioning, evaluating, setting standards for, sanctioning, etc., our thoughts, decision, and actions in light of what we *value*, positively or negatively. As such it requires a *transcendent* standpoint.

Normative thoughts are often associated with *critical* questions, and *truth* emerges as a standard for a positive answer to one, especially fundamental, normative-transcendent question: the question of *correctness* (of a given immanent thought). Roughly, and informally, the *question of truth*, or correctness, as it applies to a given immanent thought X can be expressed by:

(QT) Is the world the way the immanent thought X says it is? Do the objects that X talks about have the properties, or stand in the relations, that it attributes to them?

At issue is whether X is connected to reality in a way that warrants a positive answer to this question.¹²

¹² Elsewhere (Sher, Forthcoming 2,3) I discuss the diversity of ways in which an immanent thought can, or need, be connected to reality in order for the answer to this question to be positive.

Due to the complexity of the human cognitive situation noted above, the question of truth is one of the main engines of the cognitive-epistemic project, and *truth* is a *standard* (norm) for a positive answer to this question. When a given immanent thought satisfies this standard we say that it is *true*, or that it has the *property* of truth. When it does not, we say that it is *not true* or, in many cases, *false*.¹³ We may never be able to give a final answer to the question of truth concerning a given immanent thought, an answer that could not be questioned or challenged in return. But it is just for this reason that the question of truth is a central driving force in the search for knowledge rather than a superfluous scholastic question.

We have called immanent thoughts that attribute truth-properties to immanent thoughts “truth-thoughts”. Such thoughts, on our account, are both *normative* and *descriptive*. Inasmuch as they are concerned with the satisfaction of the truth standard, they are *normative*; inasmuch as they are concerned with the possession of the truth property, they are *descriptive*.

The fact that truth is a transcendent standard of *correctness* sets special requirements on the *transcendence* of truth-thoughts. To attribute a truth-property to a given immanent thought, to determine whether a given immanent thought is true or false, we need to transcend it not just to any standpoint from which we can see it, but to a standpoint from which we can see *both it and the world* – specifically, that part, facet, or aspect of the world (objects in the world) that it is directed at, says something about, attributes properties to. It is at this point, a point from which we measure a given immanent thought, as a content-conveying thought (rather than as a syntactic object), in relation to its subject matter in the world, that truth arises in our cognitive life.

Such a standpoint is required to be transcendent in a strong sense. It is required to have in view both the target immanent thought and that part of the world that this thought is directed at, and it must afford us a critical view of the relation between them. For that reason, in the case of truth-thoughts we are required to step *outside* the target thought and cannot stay inside it (as we can do – using Gödel numbering – if our interest is the syntax rather than the truth value of the

¹³ See discussion of bivalence in Section VI below.

target thought). This need to “step outside” means that in a sentence of the form “X is true” (“X is false”, “X is not true”, etc.), the sentence named by “X” has to stand on its own as an immanent sentence. I.e., if we take a token of “X is true” and remove “is true” from it, we are left with a name of a token *immanent sentence*. (“X is true” minus “is true” names an *immanent* sentence.) Furthermore, the transcendence inherent in truth-thoughts is *antisymmetric*: if a given truth-thought, t_1 , transcends another truth-thought, t_2 , the thought transcended by t_2 is not t_1 , and the truth predicate belonging to t_2 is not transcendent to the truth predicate belonging to t_1 .¹⁴ Now, all these constraints on the transcendence of truth-thoughts are also constraints on their immanence: For a *truth-thought* to be *appropriately immanent* its *truth predicate* must be *appropriately transcendent* to some *appropriate immanent* thought(s) in the sense indicated above. The immanence–transcendence complementarity is required to satisfy this condition.

We call the combination of principles discussed in this section the “IMMANENCE thesis”. The IMMANENCE thesis, like its constituting principles, is a material thesis. It says that *truth arises in the intersection of the immanent, transcendent, and normative modes of thought* in the way just described (be it only briefly and informally). We can sum it up by saying:

- A. The question of truth arises for all and only *immanent* thoughts.
- B. Truth is a *normative-transcendent* standard for *immanent* thoughts. (Truth and falsehood (non-truth) are *normative-transcendent* properties of *immanent* thoughts.)
- C. Truth-thoughts are *immanent, transcendent, and normative*.

The IMMANENCE thesis has significant ramifications for other principles of truth, including the *equivalence principle* and, through it, the Liar Paradox. Since the equivalence principle and

¹⁴ Note that this does not conflict with the possibility, mentioned above, of transcending philosophy in order to study it sociologically and vice versa. In the course of studying philosophy by sociology, the latter is transcendent to the former, and not vice versa, while in the course of studying sociology by philosophy the situation is reversed. So in the course of each study the transcendence relation is antisymmetry. Of course, we can transcend (the relevant parts of) both sociology and philosophy to talk about their possible interrelationships (as we are presently doing). But this view, too, involves an antisymmetric transcendence relation.

the Liar Paradox are commonly viewed as pertaining to small units of thought – sentence-like units rather than whole-theory-like units – from now on I will identify “thought” with “sentence-like thought” or, for short, “sentence”.

IV. *Ramifications for the Equivalence Principle*

The IMMANENCE thesis says that given an immanent sentence P, we transcend it to make the truth-statement that P is true, and our standard of truth says that P is true iff the world is as P says it is, i.e., on one formulation, “<P> is true iff P”. In this way, the IMMANENCE thesis supports, or gives rise to, the equivalence principle, E. But the IMMANENCE thesis, and in particular two of its subprinciples, immanence and transcendence (with significant help from normativity), also restrict E by setting certain constraints on it. Two such constraints are:

- (*Im*) For E to hold for any given sentence P, P must be a bona fide immanent sentence.
- (*Trans*) If P is a truth-sentence, then for E to hold for P, P must be appropriately transcendent. (The truth predicate of P must be appropriately transcendent.)

Accordingly, our theory of truth sanctions *only* those instances of E that satisfy the two *material* conditions, *Im* and *Trans*. Three instances of E that are sanctioned by these conditions are:

- (E₁) “Snow is white” is true iff snow is white,
- (E₂) “ ‘Snow is white’ is true” is true iff “snow is white” is true,¹⁵
- (E₃) “ ‘Snow is white’ is false ” is true iff “snow is white” is false.

E₁ says that the immanent sentence “Snow is white” satisfies our (transcendent) standard of truth, or has the (normative-transcendent) property of being true iff the world is as the sentence “Snow is white” says it is, i.e., under one formulation, iff snow is white. E₂ says that the immanent *truth-sentence* “ ‘Snow is white’ is true” satisfies our (transcendent) standard of truth, or has the (normative-transcendent) property of being true iff the world is as the immanent truth-

¹⁵ I use the convention that embedded quotation marks follow the pattern <“ ‘ , ‘ , ‘ , ...>.

sentence “‘Snow is white’ is true” says it is, i.e., under one formulation, iff “Snow is white” is true. And E_3 says that the immanent *truth-sentence* “‘Snow is white’ is false” satisfies our (transcendent) standard of truth, or has the (normative-transcendent) property of being true iff the world is as the truth-sentence “‘Snow is white’ is false” says it is, i.e., under one formulation, iff “Snow is white” is false. All these instances of the equivalence principle satisfy the requirements *Im* and *Trans*: “Snow is white”, “‘Snow is white’ is true”, and “‘Snow is white’ is false” are appropriately immanent, and the last two are appropriately transcendent.

Now, it follows from the material constraints *Im* and *Trans* that $E_P - E$ as applied to P – is an inadmissible instance of E in at least two circumstances:

- (a) P is not a bona fide immanent sentence.
- (b) P is a truth-sentence and P is not appropriately transcendent (P 's truth predicate is not appropriately transcendent).

I.e., all instances of E that satisfy either one of the material conditions (a) or (b) are blocked by our substantivist theory of truth.

Another way of expressing these constraints is by saying that *E can be instantiated only by admissible truth-bearers*, where *Im* and *Trans* are reformulated as conditions on admissible truth-bearers. Thus formulated, *Im* says that for any sentence to be an admissible truth-bearer it must be immanent, and *Trans* says that for any truth-sentence to be an admissible truth-bearer it must be appropriately transcendent (or that its truth predicate must be appropriately transcendent to it). We may say that IMMANENCE induces an update of E , or that $E \boxtimes$ IMMANENCE involves an update of E which is expressed by constraints placed on the scope of E , so that now we have:

(E) $\langle P \rangle$ is true iff P ,

where “ P ” stands for any sentence satisfying *Im* and (where applicable) *Trans*, and “ $\langle P \rangle$ ” stands for a name of this sentence.

V. *Blocking the Liar Paradox*

The Liar Paradox arises from a particular instance, or cluster of instances, of the equivalence schema. One classical representative of these instances is

(E_L) “L is false” is true iff L is false,

where “L” abbreviates “ ‘L is false’ ”. We arrive at the Liar Paradox by replacing “ ‘L is false’ ” with “L”, getting:

(LP) L is true iff L is false.

Our theory, however, blocks LP by ruling, based on the material conditions *Im* and *Trans*, that E_L is an *inadmissible instance of E* or, alternatively, that the Liar sentence,

(L) L is false.

is an *inadmissible truth-bearer*.

Before explaining why the L/E_L fails to satisfy *Im* and *Trans*, let us set the context for our explanation by a few preparatory and background comments:

- (i) In discussing whether L satisfies the *Im* and *Trans* conditions we will not appeal to any mechanical (or semi-mechanical) test. This is connected with the fact that at this stage we are not committed to any specific systematization of the immanence and transcendence principles. Systematization has many benefits and is an important goal of theorizing; but it often comes with a price: an ad hoc treatment of outliers.¹⁶
- (ii) Given our cognitive-epistemic perspective, our goal is not to rule out only (or specifically) paradoxical sentences as admissible truth-bearers; our goal is to rule out any sentence that from a cognitive perspective is not a genuine truth-bearer. This distinguishes us from theorists like Kripke who aim at ruling out only paradoxical sentences. For example, for us it might be just as important to reject

¹⁶ I should note, however, that sometimes this price is worth paying. In those cases pragmatic considerations guide us in making choices. For example, we may choose, for the sake of systematicity, to render every 1st-level predicate applicable to every singular term. This might lead us to say, on pragmatic grounds, that “The number two is laughing” is meaningful or even immanent. Due to transcendence, however, we can always move to a standpoint from which we can distinguish (and explain the differences between) genuine and conventional (or ad hoc) meaningfulness and immanence as well as genuine and conventional truth.

(T) T is true

as to reject L as an admissible truth-bearer.

- (iii) Since the Liar Paradox is commonly thought to enter theories of truth through some instance of E, in this paper we focus on E as a “gate” to the Liar. But it is worthwhile to note that from a cognitivist-epistemic perspective E is less central to theories of truth than it is often taken to be. The connection of immanent sentences to the world (through their truth conditions) is very central, but E itself offers just one particular way of representing this connection and there could in principle be alternative representations. For example, in the case of the immanent sentence “Snow is white” it is possible to represent its connection to the world (through its truth conditions) by a non-E-biconditional like:

“Snow is white” is true iff snow reflects light of all hues completely and diffusely.

And for some purposes (e.g., informativeness) this representation might be superior to the more common E-representation:

“Snow is white” is true iff snow is white.

Thus one possible way of blocking the paradox could be to remove E from the theory altogether and introduce a different equivalence principle in its place. Rejecting E altogether, however, would be ad hoc, since most instances of E are true. For that reason we prefer to approach our task by limiting the scope of E.

- (iv) Given that L is a self-referring sentence and that this trait is sometimes associated with its paradoxicality, the question naturally arises whether immanence and transcendence ban (either individually or together) all self-reference. The answer is “No”. For example, the sentence

(S) S is short,

is self-referring, yet in spite of this it is not banned by either *Im* or *Trans*. Since S is not a truth-sentence, it is not subject to *Trans*. It is subject to *Im* but it has no difficulty satisfying *Im*, since it attributes a property applicable to syntactic entities (being short) to a bona fide syntactic entity in the world, namely, to itself as a syntactic entity, or to the syntactic facet of itself (as a sentence).

- (v) Another feature that is often attributed to L is circularity. Do *Im* and/or *Trans* ban all forms of circularity? The answer, here too, is “No”. This can be seen both by considering the holistic nature of our theory, which sanctions non-vicious circularity (e.g., the kind of circularity represented by back and forth moves within Neurath’s boat), and by considering particular sentences. Consider the sentence

(A) The property of being abstract is an abstract property.

Like (S), this sentence is not subject to *Trans* and it satisfies *Im*. This it does by attributing a property to an object it is applicable to, in this case, a property. (The identity of the two properties does not change this fact.)

- (vi) To say that a given sentence is not immanent is not to reject it as a sentence or even as a meaningful sentence. Immanence identifies a subclass of sentences, namely, those that potentially partake in the cognitive project, a project that gives rise to a notion of truth (or calls for a standard of truth). But language, as we will note in Section VI below, has multiple functions, and therefore a sentence that fails to contribute to the cognitive project might yet fulfill another function that could render it meaningful (for example, the sentence “I do” uttered by a court witness while being sworn in).
- (vii) To say that a given sentence is transcendent is not necessarily to say that it belongs to a metalanguage. The object-language – metalanguage complementarity is one model (systematization) of the immanence–transcendence complementarity, but it is not its only model, as we will see in discussing Kripke below.

We are now ready to turn to the Liar sentence, L, and the corresponding E-instance, E_L. L says of itself that it is false, i.e, it attributes a truth-property to itself. As such it is a truth-sentence and therefore subject to both *Im* and *Trans*. The question is whether L satisfies these requirements. To satisfy *Trans*, the operation of truth predication in L must be appropriately transcendent to the object it is predicated of, namely L itself. I.e., in light of the conditions set on transcendence by the normativity of truth (see last paragraph of subsection of normativity) L minus “is false” must be immanent. But once we remove “is false” from L, no immanent sentence is left. So L fails to satisfy *Trans*: there is no appropriately immanent object for L to transcend.¹⁷

¹⁷ If we did not include immanence as a condition of appropriate transcendence (i.e., if we did not say that for a truth sentence to be transcendent it must be transcendent to an appropriate *immanent* sentence), then we could say that to satisfy *Im* L must fail to satisfy *Trans* and to satisfy *Trans* L must fail to satisfy *Im*.

Another way to pinpoint the blocking of L is to say that to be appropriately immanent, L must attribute its truth-property to itself as a semantic, *content conveying* sentence. But once its truth-property is removed from its content (by transcendence) it is no longer a content conveying, hence appropriately immanent, sentence.

It follows that L is an inadmissible truth-bearer according to our theory, hence E_L is an inadmissible instance of E and the Liar Paradox (in its classical form) is blocked.

We have shown that the Liar Paradox can be blocked by material principles arrived at in the course of developing a substantive theory of truth. To do that we have focused on one, albeit classical, form of the paradox. This is sufficient to make our main point, but the question naturally arises whether the same material principles are capable of blocking other forms of the paradox as well.

Let us examine two other Liar sentences, leading to other forms of the Liar Paradox. First, let us consider a type of Liar exemplified by

(7) (8) is false,

where the sentence referred to in (7) is

(8) (7) is true.

To show that (7) is not an admissible truth-bearer we note that for (7) to be an admissible truth-bearer, it must satisfy *Trans*, and this requires both that (8) be an admissible truth-bearer and that (7) be appropriately transcendent to (8). But for (8) to be an admissible truth-bearer, it has to be appropriately transcendent to (7), and this is impossible. Transcendence is an antisymmetric relation, hence (7) and (8) cannot mutually transcend each other. As in the case of L, (7) fails to satisfy *Im*.

The next “Liar sentence” we consider is only contingently a Liar sentence. Let me explain. So far we have considered sentences that give rise to the Liar Paradox due to their own semantic structure and/or the semantic structure of other sentences they are relevantly connected to. Kripke (1975) shows that the Liar Paradox can arise due to circumstances that hold in the world, so that a sentence with a given semantic structure can be paradoxical under some circumstances and non-paradoxical under others. Are Liar sentences of this kind ruled out by *Im* and *Trans*? And if they are, are they ruled out in all circumstances or only in problematic circumstances? The answer is

that they are ruled out in problematic circumstances (of the kind considered by Kripke) but not in all circumstances.

Kripke's example of a contingently-Liar sentence is:

(9) Most (i.e., a majority) of Nixon's assertions about Watergate are false. [*ibid.*: 691]¹⁸

While under most circumstances this sentence is non-paradoxical, under some circumstances it is.

Consider the following two constellations of circumstances:

Constellation of Circumstances #1:

(i) One of Nixon's assertions is

(10) Everything Jones says about Watergate is true [*ibid.*];

(ii) Aside from this sentence, "Nixon's assertions about Watergate are evenly balanced between the true and the false" (*ibid.*), and

(iii) Jones made exactly one assertion about Watergate, (9).

It is easy to see that under this constellation of circumstances (9) is paradoxical, and it is also easy to see that *Im* and *Trans* rule out (9) as an admissible truth-bearer under this constellation. Given (i)-(iii), both transcendence and immanence break down, or malfunction, in (9), as they do in (7).

Constellation of Circumstances #2:

(i) Nixon made only one assertion about Watergate.

(ii) Nixon's assertion about Watergate is:

(11) I have nothing to do with the Watergate break-in.

Clearly, under this constellation of circumstances (9) does not yield a paradox, nor is it rejected by *Im* or *Trans*.

With this we have completed the task of showing that the Liar Paradox can be blocked by substantivist theories of truth due to their material principles. It follows that philosophers need not impose Liar-motivated constraints on their theories of truth in advance, constraints that might

¹⁸ "(9)" is my own number for this sentence.

unnecessarily limit their options in constructing their theories. They may wait until a later stage in the construction of their theory to decide which constraints (among all the alternatives offered by different solutions to the paradox) are most suitable for their theory, *if* the need for such constraints arises at all.

Our next task is to explore how changing our attitude toward the Liar Paradox in this way affects our perspective on existent solutions to the paradox and on some of the factors commonly associated with these solutions.

VI. Tarski, Kripke, Natural Language, and Bivalence

It is quite common to view Tarski's (1933) hierarchical solution to the Liar Paradox as an unnecessarily radical, ad hoc solution and contrast it with other, less radical and less ad hoc solutions like Kripke's (1975). Our discussion of truth, transcendence, and the Liar puts the two types of solution in a new perspective: both solutions are based on a *non-ad-hoc* systematization of the immanence–transcendence complementarity, Tarski's on an *external-hierarchy* systematization and Kripke's on an *internal-hierarchy* systematization, and each systematization has its own advantages and disadvantages.

Tarski's External Hierarchy as a Systematization of the Immanence–Transcendence Complementarity. In the Tarski systematization the immanence–transcendence complementarity is represented by an object-language–metalanguage hierarchy. We start with a so called object language, L_0 , which is a purely *immanent* language. Sentences of this language are directed at something external to themselves, something in the world (broadly construed). The language has no resources for transcendence, e.g., no direct resources for naming its own sentences and no semantic predicates (or tools for creating them). This object language serves as a basis for the construction of a hierarchy of languages and is the lowest element in this hierarchy. The next language is a so called metalanguage, ML_0 or L_1 , and we arrive at it by *transcending* L_0 . L_1 is fully equipped for (truth-) transcendence: it has resources for talking about L_0 , resources for

talking about that part or aspect of the world that L_0 is directed at, and resources for setting a standard of truth for sentences of L_0 . Furthermore, L_1 is itself an immanent language, and as such has its own metalanguage, L_2 ($=ML_1 = MML_0$) to which it stands as an object language. Each metalanguage has a truth predicate applicable to its predecessor in the hierarchy, and as a result, truth itself is represented by a hierarchy of predicates: T_1, T_2, T_3, \dots , which apply to sentences of L_0, L_1, L_2, \dots , respectively.

As a systematization of the immanence–transcendence complementarity Tarski’s hierarchy has both advantages and disadvantages. Among its advantages are a clear, sharp, unambiguous, and formally rigorous structure, as well as one that has proved extremely fruitful in multiple fields of knowledge. Let me briefly reflect on its fruitfulness. It is a remarkable fact that while Tarski used the object-language–metalanguage duality as a formal solution to a formal problem concerning the largely material notion of truth, this duality, partly under his influence, has been thoroughly integrated into our conception of knowledge and has proved extremely profitable in a number of fields, in particular metalogic and metamathematics. My present suggestion is that one of the reasons this duality has been a major force in advancing human knowledge is that it captures the cognitively fundamental and highly fruitful complementarity of immanence and transcendence.

Among the weaknesses of the Tarskian hierarchy, from our perspective, are a rigid and inflexible structure, as reflected in, e.g., the exclusion of many self-referring sentences satisfying *Im* and (if applicable) *Trans*. Its rigidity is also reflected in some artificial aspects of its treatment of truth, for example, the existence of multiple truth predicates, each limited to a particular language in the hierarchy, rather than a single predicate, common to the entire hierarchy. Another weakness is the narrow range of languages that Tarski’s hierarchy encompasses, namely, its limitation to so called “formalized languages of the deductive sciences”. These languages by no means exhaust the full range of cognitively efficacious languages licensed by the (pre-systematic) principles of immanence and transcendence.

A different model of the immanence–transcendence complementarity is suggested by Kripke’s solution to the Liar Paradox.

Kripke’s Internal Hierarchy as a Systematization of the Immanence–Transcendence Complementarity. In Kripke’s model we have only one language and only one truth predicate. The idea of immanence is represented by the concept of *groundedness* and the idea of transcendence by a *stage by stage* determination of the extension (and counterextension) of the truth predicate. Kripke informally describes the intuition underlying his concept of groundedness as follows:

It has long been recognized that some of the intuitive trouble with Liar sentences is shared with such sentences as [T¹⁹] which, though not paradoxical, yield no determinate truth conditions. ... In general, if a sentence ... asserts that (all, some, most, etc.) of the sentences of a certain class C are true, its truth value can be ascertained if the truth values of the sentences in the class C are ascertained. If some of these sentences themselves involve the notion of truth, their truth value in turn must be ascertained by looking at *other* sentences, and so on. If ultimately this process terminates in sentences not mentioning the concept of truth, so that the truth value of the original statement can be ascertained, we call the original sentence *grounded*; otherwise, ungrounded. ... Sentences such as [T], though not paradoxical, are ungrounded. [Kripke 1975: 693-4]

The grounded sentences in Kripke’s system are all immanent in our sense, and most immanent sentences in our sense are included in the class of Kripke’s grounded sentences.²⁰

The idea of transcendence is represented by Kripke’s conception of the (single) truth predicate as defined *in stages*, or its extension as determined in stages. In *Stage 1* the truth predicate, *Tr*, is assigned an extension (and a counterextension)²¹ in the set of all

¹⁹ I.e., the sentence we called “T” in Section V above.

²⁰ For important exceptions, see below. I should note that Kripke, like us, distinguishes between meaningful and immanent (for him, grounded) sentences. (See *ibid.*: 699-700.)

²¹ The counterextension of *Tr* is the set of all *false* sentences, leaving truth-wise indeterminate sentences outside both extensions. (Kripke’s system is based on a trivalent logic. (See discussion of bivalence below.))

grounded/immanent sentences that contain no truth predicate (or semantic predicates more generally). In *Stage 2* the extension (and counterextension) of *Tr* is extended to all grounded/immanent sentences whose truth value is determined in Stage 1. And so on. In this way, the assignment of a truth value to each immanent sentence in Kripke's system is made from a standpoint transcendent to it, albeit internal to the (single) Kripkean language, and Kripke's system can be viewed as representing the immanence–transcendence complementarity by an *internal hierarchy*. Things change, however, once we get beyond the “‘minimal’ or ‘smallest’ fixed point” (*Ibid.*: 705), i.e., beyond the point in which the hierarchy is limited to grounded/immanent sentences. Beyond this point Kripke deals with sentences like *T*, which are neither grounded nor paradoxical, and shows how a conventional assignment of truth values to such sentences can proceed. Since these sentences are non-immanent (ungrounded), they do not fall under our immanence and transcendence principles, and as such are not admissible truth-bearers according to our theory.²²

One important feature of Kripke's system is its allowing a sentence to “seek its own level” (*ibid.*: 696) in the internal hierarchy. Whether or not a given sentence has a place in the internal hierarchy, and if it does, what this place is, Kripke points out, may be determined by empirical circumstances surrounding the sentence. It is in this context that he introduces the sentence numbered (9) in the present paper, which, he rightly says, is an admissible truth-bearer in some circumstances, inadmissible (and paradoxical) in others.

Among the advantages of Kripke's hierarchy over Tarski's from our perspective are its considerable flexibility and highly dynamic nature. These are reflected in its ability to recognize the admissibility of sentences that are excluded from Tarski's hierarchy as truth-bearers (e.g., some self-referring sentences satisfying *Im* and (where applicable) *Trans* that are banned from

²² The sentences “*T* is not paradoxical” and “Technically, *T* can be assigned a truth value in a consistent system”, however, are immanent according to our theory, and therefore the upper echelon of Kripke's system does make a significant contribution to understanding truth, according to our theory, in spite of dealing with empty (inadmissible) truth-bearers.

Tarski's hierarchy) as well as in its ability to account for the role of empirical circumstances in the admissibility and place in the hierarchy of sentences like (9). Importantly, its flexibility saves it from the need to artificially multiply truth predicates as in Tarski's hierarchy.²³

But Kripke's system has significant disadvantages as well. In particular, its internal hierarchy is too confined to encompass the whole Tarskian hierarchy, leaving many metalinguistic truth-bearers unaccounted for. Kripke himself is fully aware of the indispensability of an (external) hierarchy like Tarski's:

It seems likely that many who have worked on ... [non-Tarskian] approach[es] to the semantic paradoxes have hoped for a universal language, one in which everything that can be stated at all can be expressed. ... Now the languages of the present approach contain their own truth predicates and even their own satisfaction predicates, and thus to this extent the hope has been realized. Nevertheless the present approach certainly does not claim to give a universal language, and I doubt that such a goal can be achieved. First, the induction defining the minimal fixed point is carried out in a set-theoretic metalanguage, not in the object language itself. Second, there are assertions we can make about the object language which we cannot make in the object language. For example, Liar sentences are *not true* in the object language, in the sense that the inductive process never makes them true; but we are precluded from saying this in the object language by our interpretation of negation and the truth predicate. If we think of the minimal fixed point ... as giving a model of natural language, then the sense in which we can say, in natural language, that a Liar sentence is not true must be thought of as associated with some later stage in the development of natural language, one in which speakers reflect on the generation process leading to the minimal fixed point. It is not itself a part of that process. The necessity to ascend to a metalanguage may be one of the weaknesses of the present theory. The ghost of the Tarski hierarchy is still with us. [*Ibid.*: 714]

Natural Language and the Cognitive Perspective on Truth. Many of the objections to Tarski's hierarchy in the philosophical literature have to do with natural language. In a way, Tarski himself invited these objections. In setting the ground for his theory of truth he described his goal as constructing a philosophical definition of the concept of truth, and his explanation

²³ Another purported advantage of Kripke's system is its ability to account for the behavior of the truth predicate of natural language. I will briefly discuss the relevance of natural language to the theory of truth below.

suggested that it is “the meaning of the term ‘true sentence’ in colloquial language” (Tarski 1933: 152) that such a definition aims at capturing. Yet, Tarski emphasized that a “thorough analysis of the meaning current in everyday life of the term ‘true’ is not intended here” (*ibid.*: 153), and he proceeded to question the very consistency of natural language on the ground that it leads to paradox: Due to the occurrence of the Liar Paradox in natural language, “*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 very questionable, and consequently the same doubt attaches to the possibility of constructing a correct definition of this expression*” (*ibid.*: 165). Finally, he formulated his goal as that of defining truth for a narrow class of highly specialized artificial languages: the “formalized languages” of “the *deductive sciences*” (*ibid.*: 166). It is not surprising, therefore, that many philosophers questioned Tarski’s approach to the construction of a theory of truth, starting with his sweeping conclusion about the inconsistency of natural language and ending with his narrow focus on artificial logical languages.²⁴

Leaving the controversy over Tarski’s approach to truth aside, let me briefly explain how natural language is viewed from the perspective of our theory – both from its broader *substantivist* perspective and from its narrower, though still very broad, *cognitivist* perspective. From a substantivist perspective the diverse, multifaceted, and sometimes conflicting uses of “truth” and its cognates in natural language(s) reflect the extraordinarily breadth and diversity of truth, and an account of the breadth and diversity of truth and their manifestations in natural language is one of the tasks of a substantive theory (family of theories) of truth. But it is not the only task, or necessarily the most important philosophical task, of the theory of truth.

²⁴ My own view, which I have expressed elsewhere (Sher 1999b, 2004), is that what Tarski’s theory *actually* does is provide a model of the *logical factor* in truth, and in particular, the role logical structure plays in the truth conditions of sentences. This justifies Tarski’s choice of *logically* formalized languages to develop his definition, and it explains why his definition of *truth* had such fruitful applications in *metalogue*.

Indeed, there are several complicating factors in focusing the philosophical study of truth on natural language. Among these are (i) the fact that natural language is a *natural* phenomenon, and (ii) the fact that natural language is a *multipurpose* tool. These facts suggest that the use of “truth” in natural language is shaped by multiple, sometimes conflicting and often accidental determinants, and that as a result natural language is less than an optimal source of understanding the role of truth in specific human endeavors, including cognition or knowledge. This is the point at which we, given our interests, distance ourselves from natural language; but it is also the point at which we distance ourselves from Tarski’s formalized languages. We are investigating truth from a specific, yet relatively broad perspective, the cognitivist-epistemic perspective, and this means that on the one hand we cannot limit ourselves to the behavior of truth in formalized languages, but on the other hand we are not concerned with all its manifestations in natural language. Moreover, our interest in truth is partly normative. This means that as far as language is concerned we are interested in a model of language that is recognizably a model of human language, but one that is geared toward effective cognition of the world rather than toward other tasks. Thus, from our perspective, a study of the behavior of the term “true” in natural language is valuable in understanding some aspects of truth, but for understanding the role of truth in human knowledge and cognition natural language is secondary.

The Immanence–Transcendence Complementarity and Bivalence. It is common to contrast Tarski’s and Kripke’s hierarchies not just along the formal-language vs. natural-language dimension, but also along the bivalence vs. trivalence (or more generally non-bivalence) dimension. From our perspective, these contrasts are only partly warranted. In particular, the viability of Tarski’s and Kripke’s hierarchies as models of the cognitive immanence–transcendence complementarity has little to do with these contrasts. Focusing on the bivalence–trivalence dichotomy, let me emphasize, first, that the immanence–transcendence complementarity is perfectly compatible both with classical logic and with nonclassical logic, both with classical set theory and in nonclassical set theory. There is nothing in our account of

immanence or transcendence to favor one logic (or set theory) over the other. If the world, or some aspects of the world, exhibit a tripartite rather than a bipartite property structures, then immanence will be better represented by a nonclassical logic (and/or a nonclassical set theory) than by a classical one. The account of the formal structure of properties will have a trivalent rather than a bivalent operation of complementation, so that given a domain D and a property P , complementation will partition D into three rather than two regions. Transcendence would then require a standpoint from which we can view the world in its tripartite structure, and the determination of truth values would be made based on this view of the world, hence involve three rather than two truth values. There is thus no intrinsic sense in which Tarski's notions of object language and metalanguage, or our notions of immanence and transcendence, or our material conditions of *Im* and *Trans*, are tied up with, or limited to, classical logic or bivalence. Significant structural differences will still exist between Kripke-like and Tarski-like models of the immanence–transcendence complementarity, but the differences between formalized and natural language on the one hand and bivalence vs. trivalence on the other are not central to them.

References

Dyson, F. 1988. *Infinite in All Directions*. Harper & Row.

Frege, G. 1892. "On Sense & Nominatum". Tr. H. Feigl. *Readings in Philosophical Analysis*. Eds. H. Feigl & W. Sellars. 1949. New York: Appelton-Century-Crofts, 1949. Pp. 85-102.

Gödel, K. 1931. "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems". *Kurt Gödel: Collected Works*, Volume 1. Eds. S. Feferman et al. Oxford, 1986. Pp. 145-95.

James, W. 1890. "The Stream of Thought". *The Principles of Psychology*, Volume 1. New York: Cosimo, 2007. Pp. 224-290.

Kripke, S. 1975. "Outline of a Theory of Truth". *Journal of Philosophy* 72: 690-716.

Quine, W.V. 1970/86. *Philosophy of Logic*. McGraw Hill.

----. 1981. "Things and Their Place in Theories". In Quine (1981a): 1-23.

- . 1986. "Reply to Harold N. Lee". In *The Philosophy of W.V. Quine*. Eds. L.E. Hahn & P.A. Schilpp. La Salle: Open Court. Pp. 315-8.
- . 1995. "Reactions". In *On Quine: New Essays*. Eds. P. Leonardi & M. Santambrogio. Cambridge. Pp. 347-61.
- Sher, G. 1999a. "On the Possibility of a Substantive Theory of Truth". *Synthese* 117: 133-72.
- . 1999b. "What Is Tarski's Theory of Truth?" *Topoi* 18 (1999): 149-166.
- . 2004. "In Search of a Substantive Theory of Truth". *Journal of Philosophy* 101:5-36.
- . 2010. "Epistemic Friction: Reflections on Knowledge, Truth, and Logic". *Erkenntnis* 72: 151-76.
- . 2012. "Truth & Knowledge in Logic & Mathematics". *The Logica Yearbook 2011*. Eds. M. Peliš & V. Punčochář. London: College Publications, King's College. Pp. 289-304.
- . Forthcoming 1. "The Foundational Problem of Logic". *The Bulletin of Symbolic Logic*.
- . Forthcoming 2. "Forms of Correspondence: The Intricate Route from Thought to Reality". *Truth & Pluralism: Current Debates*. Eds. N.J.L.L. Pedersen & C.D. Wright. Oxford University Press: 157-179.
- . Forthcoming 3. "Truth as Composite Correspondence". *Unifying the Philosophy of Truth*. Eds. T. Achourioti, H. Galinon, H. Fujimoto, and J. Martinez-Fernández. Dordrecht: Springer.
- Siewert, C. 2006. "Consciousness & Intentionality". *Stanford Encyclopedia of Philosophy*.
- Tarski, A. 1933. "The Concept of Truth in Formalized Languages". *Logic, Semantics, Metamathematics*. Tr. J.H. Woodger. Indianapolis: Hackett (1956/83): 152-278.
- Wittgenstein, L. 1921. *Tractatus Logico-Philosophicus*. Trs. Pears & McGuinness. Routledge & Kegan Paul, 1961.