

LESSONS ON TRUTH FROM KANT*

GILA SHER

University of California, San Diego

Kant is known for having said relatively little about truth in *Critique of Pure Reason* (1781/7), and most commentators have followed suit. Many (including Bennett 1966; Strawson 1966; Wolff 1973; Hossenfelder 1978; Allison 1983; Guyer 1987; Longueness 1993; Gardner 1999; and others) have no entry for “truth” in their index, and others have only few references for this term.¹ Nevertheless, there are important lessons to be learned from Kant about truth, lessons that apply to the contemporary debate on the nature and structure of truth and its theory.² In this paper I would like to suggest two such lessons, one negative, the other positive. Both lessons concern the structure of a *substantive* theory of truth as contrasted with a *deflationist* theory;³ the second lesson concerns the *correspondence* theory of truth as well.⁴ In deriving these lessons I will largely focus on Kant’s first *Critique*, and in particular on its first half (up to the Transcendental Dialectic).⁵

I. A Negative Lesson from Kant on Truth

In the third section of the Introduction to Transcendental Logic Kant distinguishes two distinct questions of truth: a question he regards as

*Earlier versions of this paper were presented in my graduate seminar “Truth in Kant”, a meeting of the UCSD History Round Table, and talks at the Southern California Epistemology group meeting at the Pacific APA and the Seventh Annual Analytic Philosophy Symposium at the University of Texas at Austin. I would like to thank the participants for valuable feedback and discussion. On a personal note, I would like to thank Bob Hanna, Michael Hardimon, Michael Lynch, Michael Oberst, Carl Posy, Sam Rickless, Peter Sher, Clinton Tolley, Eric Watkins, and Cory Wright for very helpful comments.

¹ E.g., Cohen (1907). There are a few notable exceptions—e.g., the authors mentioned in fn. 6—but they are a minority.

² Hanna (2000) and Rosenkoetter (2009) also see Kant’s view of truth as relevant to the contemporary debate.

³ For substantive/substantivist theories of and approaches to truth, see, e.g., Wright (1992), Gupta (1993), David (1994), Lynch (1998, 2009), and Sher (1999, 2004, 2016a,b). For deflationist theories of and approaches to truth, see, e.g., Ramsey (1927), Leeds (1978), Field (1986, 1994), Horwich (1990/8), and Williams (1999).

⁴ I will address the view that Kant’s conception of truth is coherentist or epistemic-pragmatist in Part II, section 3.2-3.

⁵ References to *Critique of Pure Reason* include first and second edition original paging. Citations are based either on the Guyer and Wood translation or on the Kemp Smith translation. In citations from the latter I change the translation of “Erkenntnis” from “knowledge” to “cognition”.

legitimate and a question he regards as illegitimate.⁶ The legitimate question is “[What is the] **nominal definition** of truth[?]”⁷ (A58/B82) or, using a more literal translation, “[What is the] *name-clarification* of truth”[?]. This question Kant contrasts with the question “[W]hat is the general and sure **criterion** of the truth of **any and every** cognition[?]” (ibid.). The answer to the first question, Kant says, is that “truth... is the **agreement of cognition with its object**” (ibid.). But asking the second question, he says, is “absurd” (ibid.). This question subjects those who raise it to the embarrassment of “sophism”, “ignorance”, and “emptiness”, “call[ing] for an answer where none is required” (A57-8/B82). Sarcastically, he likens those who ask this question and those who pay attention to them to “one man milking a he-goat and the other holding a sieve underneath” (A58/B8-3).

Now, a deflationist might be tempted to identify the second question with a *substantivist* approach to truth, interpreting Kant’s rejection of this question as a sign that Kant is a *deflationist* about truth. But this conclusion is too hasty.⁸ Before explaining why, however, let me briefly indicate what I mean by “deflationist” and “substantivist” approaches to truth in this paper.⁹

By a *deflationist approach to truth* I understand here the view called “minimalism” by Horwich (1990/8) and, in particular, the view expressed in the following citations from his book:

[T]ruth is entirely captured by the... triviality [“that each proposition specifies its own condition for being true (e.g. the proposition *that snow is white* is true if and only if *snow is white*)”], so that in fact **nothing could be more mundane and less puzzling than the concept of truth.**[Horwich 1990: xi]

[It is a] **misconception... that truth has some hidden structure awaiting our discovery...** [U]nlike most other predicates, **‘is true’ should not be expected to participate in some deep theory of that**

⁶ For earlier works addressing Kant’s discussion of truth in this section see, e.g., Prauss (1969), Van Cleve (1999), Sher (1999, 2004, 2016a), Hanna (2000), Rosenkoetter (2009), and Vanzo (2010). Different authors vary in their focus. Prauss, for example, focuses on the way Kant’s discussion of truth in this section fits in with the rest of the Introduction to Transcendental Logic. Unlike Prauss and some of the other commentators, my own focus lies in the actual line of reasoning presented by Kant’s main argument in this section.

⁷ Throughout the paper I use boldface within quotations for my own emphases. In citations from the Guyer & Wood translation of the *Critique* I replace their boldface by italics.

⁸ Hanna (2000, 2004/13) and Rosenkoetter (2009) also consider the relation between Kant’s discussion of truth and the deflationist view (or what is in effect a deflationist view). Like me, they reject the deflationist reading of Kant, though from a somewhat different perspective and without drawing the lesson that I will draw below.

⁹ Some philosophers use “inflationist” or “inflationary” where I use “substantive” or “substantivist”. But what I mean by this term (see below) is better captured by the latter choice of words.

to which it refers—a theory that goes beyond a specification of what the word means. [*Ibid.*: 2]

By a *substantivist view of truth* I mean the opposite view, namely, the view that truth is *not* “captured by” mere “trivialit[ies]”, that it does have a “structure awaiting our discovery”, and that the theory of truth is, or at least ought to be, a “deep theory of that to which it refers—a theory that goes [far] beyond a specification of what the word means”. This “contrastive” characterization, however, tells us very little about what a substantive theory of truth is or should be like. Indeed, it is compatible with multiple positive conceptions of a substantive theory of truth, some of which (as we will see below) are unviable. It is thus important to distinguish between viable and unviable conceptions of a substantive theory of truth, and it is here that Kant’s discussion is instructive.

The conception of a substantive theory of truth criticized by Kant I will call “the *critical* conception”. According to this conception, the task of a substantive theory of truth is to provide a “**general and sure criterion of the truth of any and every cognition**”. Kant’s argument against this conception *ibid.* proceeds as follow:

If truth consists in the agreement of a cognition with its object, then this object must thereby be distinguished from others; for a cognition is false if it does not agree with the [unique] object to which it is related. . . Now **a general criterion of truth would be that which was valid of all cognitions without any distinction among their objects. But it is clear that since with such a criterion one abstracts from all content of cognition (relation to its [unique] object), yet truth concerns precisely this content, it would be completely impossible and absurd to ask for a mark of the truth of this content of cognition, and thus it is clear that a sufficient and yet at the same time general sign [Kennzeichen] of truth cannot possibly be provided. . . [O]ne must therefore say that no general sign of the truth of the [content] of cognition can be demanded, because it is self-contradictory.**¹⁰[A58-9/B83]

More recently, a similar criticism was presented in Blackburn (1984), although in different terms:

[C]ompare ‘is true’ . . . with a genuine target of philosophical analysis. . . We know *individually* what makes [‘is true’] applicable to the judgements or sentences of an understood language. . . The reason the. . . sentence [‘Penguins waddle’] deserves the predicate is that penguins waddle, and the reason why the judgement that snow is white deserves the predicate is that snow *is* white. But these reasons are entirely different. There is no single account, or even little family of accounts, in virtue of which each deserves the predicate, for

¹⁰ A similar argument appears in Kant (1770’s–1800: 558).

deciding whether penguins waddle has nothing much in common with deciding whether snow is white. There are *as* many different things to do, to decide whether the predicate applies, as there are judgements to make. So how *can* there be a unified, common account of the “property” which these quite different decision procedures supposedly determine?[Blackburn 1984: 230]

It is, however, Kant’s more theoretical characterization of, and argument against, the criterial theory of truth, that I am interested in here.

Focusing on the actual content of Kant’s argument, we can formulate a Kantian characterization of a *criterial* theory of truth as follows:

A Kantian Characterization of a Criterial Theory of Truth:

*A criterial theory of truth provides a general and sufficient criterion of truth, where such a criterion is understood as determining, all by itself, with respect to each and every cognition (judgment, thought, belief, sentence of our language, etc.), whether it is true or false by providing a precise sign of its truth-value.*¹¹

Based on this characterization, the Kantian argument against a criterial theory of truth can be summed up as follows:

- 1 “[T]ruth consists in the agreement of a cognition with its [unique] object”.
- 2 “[A] general criterion of truth would be... valid of all cognitions without any distinction among their objects”.
- 3 “[S]uch a criterion [being general]... abstracts from all content of cognition (relation to its [unique] object)”.
- 4 “[But] truth concerns precisely this content”.
- 5 “[Hence] it would be completely impossible and absurd to ask for a [general] mark of the truth of [the] content of cognition”.
- 6 [Conclusion:] “[N]o general sign of the truth of the [content] of cognition can be demanded, because it is self-contradictory”.

A more general conclusion is:

- 7 *A criterial theory of truth is unviable.*

And a still more general conclusion is:

- 8 *A substantive theory of truth is unviable.*

¹¹ For the purpose of the present paper there is no need to worry about the identity of truth-bearers. Following Kant, I will usually speak about cognitions and judgments, but what I say applies to other truth-bearers as well. (For views and discussions of Kant’s notion of a truth-bearer see Hanna 2000; Vanzo 2010; and Tolley Forthcoming.)

But the last conclusion is not necessarily the right conclusion to draw from Kant's argument. An alternative conclusion is:

9 *A substantive theory of truth should not be conceived as a criterial theory.*

This, in my view, is the right lesson for contemporary theorists of truth to derive from Kant's argument:

Negative Lesson on Truth from Kant. Don't think of a substantive theory of truth as a criterial theory of truth in Kant's sense. Thinking of a substantive theory of truth in this way is self-defeating.

It might be worthwhile to note that some commentators (e.g., Hanna 2000: 227, 2004/13 and Vanzo 2010: 166) identify Kant's idea of a "criterion" with his idea of a "real definition".

Kant explains the notion of real definition as follows:

[A] real definition... does not merely substitute for the name of a thing other more intelligible words, but contains a clear property by which the defined *object* can always be known with certainty, and which makes the explained concept serviceable in application. Real explanation would be that which makes clear not only the concept but also its *objective reality*. Mathematical explanations which present the object in intuition, in conformity with the concept, are of this latter kind.¹²[A241-2, fn. a.]

But whether Kant uses "criterion" as a synonym of "real definition" in the argument we have discussed is of secondary importance for the lesson we drew from that argument. Our lesson is based on the specific content of the argument, and this content does not change whether we identify "criterion of truth" with "real definition of truth" or not.

Our negative lesson on truth from Kant raises the question whether there is a viable alternative to the criterial conception of a substantive theory of truth. To answer this question, I will proceed as follows: First I will generalize the negative lesson from Kant on truth to a lesson about a wide range of philosophical theories, including the theory of language, ontology, epistemology, etc. Then I will point to Kant's epistemic theory as an example of a non-criterial yet substantive philosophical theory, and I will use this theory to delineate a general, informal, characterization of substantive philosophical theories. This general characterization will be applicable to the theory of truth, along with other philosophical theories. Next, I will note that elsewhere in the *Critique* Kant recognizes a tension between generality

¹² See also Kant (1770's-1800): 215-18, 360-2, 493, and 634.

and particularity in all knowledge, and I will briefly describe his recommended solution to this tension in the general case. This recommendation, I will suggest, is applicable to philosophical theories as well, including the theory of truth. Finally, in the second half of the paper I will draw a more specific positive lesson for a substantive theory of truth from Kant, relying on certain aspects of the *Critique* that we have not discussed yet.

To generalize our negative lesson from Kant beyond truth, let us first observe that both the debate over deflationism versus substantivism and the idea of a criterial theory are applicable to a wide array of philosophical fields. Take, for example, the field of knowledge (epistemology). A deflationist about knowledge will say that all a theory of knowledge can and ought to do is to provide a minimalist definition of knowledge, while a substantivist will say that it can and ought to do far more than that. And one way to conceive of a substantive theory of knowledge is as a *criterial* theory, namely, as a theory whose task is to provide a *general and sufficient criterion* for knowledge—one that determines, all by itself, with respect to any belief or cognition, whether, or in virtue of what, it has or lacks the status of knowledge. Now, it is easy to see that if we view a substantive theory of knowledge as a criterial theory in this sense, it will be subject to essentially the same Kantian criticism as the criterial theory of truth. That is, the Kantian argument can easily be transformed to an argument against the viability of a criterial theory of knowledge.

Adopting, for the sake of the present exercise, the traditional view of knowledge as justified true belief (cognition or judgment, using Kantian terminology), the Kantian argument against a criterial theory of knowledge would take some such form as:

- 1 *For a cognition to be knowledge, (the truth of) its precise content has to be justified.*
- 2 “[A] *general criterion of [knowledge] would be... valid of all cognitions without any distinction among their [content]”.*
- 3 “[S]uch a criterion [being general]... abstracts from all content of cognition”.
- 4 “[But knowledge] *concerns precisely this content”.*
- 5 “[Hence] it would be completely impossible and absurd to ask for a [general] mark of the [knowledge] of [the] content of cognition”.
- 6 [Conclusion:] “[N]o general sign of the [knowledge] of the [content] of cognition can be demanded, because it is self-contradictory”.

Two more general conclusions, formulated in contemporary terms, are:

- 7 *A criterial theory of knowledge is unviable.*
- 8 *A substantive theory of knowledge is unviable.*

But, as in the case of truth, there is an alternative to the last conclusion, namely:

- 9 *A viable substantive theory of knowledge will not be a criterial theory.*

Generalizing, let X be a theory of a philosophical subject matter such as truth, knowledge, meaning, reference, or objects.

Generalized Negative Lesson from Kant: *Don't conclude from the fact that a criterial theory of X is unviable that a substantive theory of X is unviable. A substantive theory of X need not be (and in many cases should not be) a criterial theory.*

The challenge is to identify a non-criterial conception of a substantive theory of X. Here, however, we do have evidence that non-criterial substantive theories of some X are possible. In particular, Kant's theory of knowledge is such an example. Clearly his theory of knowledge is neither deflationist nor criterial, but (regardless of its ultimate success) it is a bona fide substantive theory, indeed a paradigm of a substantive philosophical theory. Based on this paradigm we may arrive at a

General Characterization of a Substantive Philosophical Theory: *A substantive philosophical theory of X provides a rich, significant, fundamental, explanatory, systematic, and rigorous account of X. Such a theory is not content with trivial (superficial) principles or accounts of principles. It seeks to investigate in depth central philosophical questions concerning X, solve significant philosophical problems concerning it, discover new principles governing it, and devise new methods for investigating/solving/discovering such questions/problems/principles.*

Now, this characterization is not proposed as a definition of, or a necessary and sufficient condition for, a substantive philosophical theory. It is an attempt to capture, in a common-sensical manner, a vision of a substantive philosophical theory realized by Kant in *Critique of Pure Reason* and relevant to us today. And it is applicable to the theory of truth along with other philosophical theories.

Returning to the tension between generality and particularity, we should note that Kant believes a *moderate* tension between generality and particularity (unity and diversity), the source of which is human reason, arises in *all* fields, and this moderate tension (unlike the *severe* tension arising in *criterial* theories) is *resolvable*. Thus, in the section *The Regulative Employment of the Ideas of Pure Reason*, he says that

“reason [in general]... exhibits a twofold, self-conflicting interest, on the one hand interest in... universality..., and on the other hand in... determinateness... in respect of the multiplicity” (A654/B682). Other terms he uses for this tension are “homogeneity” (*ibid.*), “unity” (A655/B683), and “identity” (A654/B682) vs. “heterogeneity” (A655/B683), “variety” (A657/B685), and “manifoldness and diversity” (A654/B682). Both generality and diversity are important, but the tension between them is resolvable. The resolution consists in *balancing* the two interests. Kant thus affirms a trio of principles: a principle of “*homogeneity*”, a principle of “*variety*”, and a principle of “*affinity*” connecting the first two (A657/B685). In our time, the tension pointed out by Kant has been widely discussed with respect to scientific theories under the rubric of “unity vs. disunity”. The Kantian resolution is eloquently expressed by Dyson who, speaking about science, says that “every science needs for its healthy growth a creative balance between unifiers and disunifiers” (1988: 47). In my view, this solution applies to philosophy as well, including its substantive theories.

Having made these general points, let us proceed to a more specific lesson (or cluster of lessons) on truth from Kant, specifically, a lesson (lessons) about the form a substantive, non-criterial, and balanced theory of truth can and should take.

II. A Positive Lesson from Kant on Truth

The key to drawing a positive lesson from Kant on truth—both truth itself and its theory—lies, in my view, in the central, if largely implicit, role that truth plays in Kant’s epistemic theory. This theory contains, I will argue, the rudiments of a substantive yet non-criterial theory of truth, a theory that is epistemic in one sense, though not in the sense in which “an epistemic theory of truth” is commonly used today. Today, by “an epistemic theory of truth” we usually understand a theory that reduces “truth” to some *narrowly* epistemic notion, such as “justification”, “evidence”, “acceptance at the ideal end of the pursuit of knowledge”, and so on. I will call an epistemic theory of truth of this kind “a *narrowly* epistemic theory of truth”. In contrast, the theory of truth tacitly embedded in Kant’s epistemology is *broadly* epistemic. Such a theory approaches truth from an epistemic *perspective*, construing it as *central* to knowledge yet not as narrowly epistemic, and using its role in knowledge as a key to understanding truth itself. The main point is that (i) knowledge requires more than just narrowly epistemic standards, (ii) truth is one of the central non-narrowly-epistemic standards it requires, (iii) one of the most important roles of truth is its role in knowledge, and (iv) therefore, understanding the role of truth in knowledge is a key (though not the only key) to understanding the nature and structure of truth.

On this construal, truth is primarily a *standard* for cognitions, so that a cognition has the property of truth if and only if it satisfies this standard.

1. Kant's Epistemic Project¹³

To understand the role of truth in Kant's theory, it will be helpful to have a general understanding of his project and of the overall structure of his argument in the *Critique*.

Acquiring such an understanding, however, is not a simple task, and this task is further complicated by the existence of a vast number of conflicting interpretations of the *Critique*. Nevertheless, an overall understanding of the Kantian project is necessary for a fruitful engagement with Kant as a player in the contemporary debate. To arrive at such an understanding I will identify Kant's goals in a way that transcends their historical context while attending to this context when describing Kant's specific treatment of the issues involved. In view of the enormous diversity of opinions among Kant scholars, I will minimize reference to secondary literature. Instead, I will rely on the Kantian text, limiting my appeal to secondary sources only when this is essential for the task at hand. What I will offer is best viewed as a proposal for understanding the overall structure of Kant's project in a way that is at once sympathetic to him and facilitates a philosophical dialogue with him in a contemporary context.¹⁴

The central goal of Kant's project, on the present proposal, is to establish the possibility of human knowledge in light of (i) skeptical challenges (e.g., Hume's) and (ii) problems arising from earlier attempts to establish its possibility (e.g., Leibniz's). It is important to emphasize that the knowledge Kant seeks to establish is knowledge of the *world*, including *scientific* knowledge of the world. In particular, Kant seeks to establish the possibility of powerful scientific knowledge, the kind of knowledge that was challenged by Humean skepticism, namely, knowledge of causal relations and laws of nature. To achieve this goal, Kant develops a new philosophical methodology (Copernican revolution), one that assigns a crucial role to the *structure of human cognition* in establishing the possibility of *bona fide human*

¹³ In this paper I share the majority view that the main project of the *Critique* is epistemic. Nevertheless, much of what I say in the rest of the paper can be adjusted to other construals of Kant's project (e.g., a semantic construal, such as Posy's 2000 or Hanna's 2001).

¹⁴ I should emphasize, though, what I have already hinted at in the last footnote, namely, that my main lessons from Kant are not conditional on the overview proposed in this section. The overview will serve primarily as a compass, guiding us in our contemporary investigation of Kant's approach to some of the fundamental questions pertinent to his implicit conception of truth.

knowledge. This methodology enables him to construct a new argument—indeed, a new kind of argument—for the possibility of knowledge of the world, and in particular, knowledge of causal relations and laws of nature.

We may view Kant's argument for the possibility of human knowledge as proceeding in the following way:

Starting with the modest assumption that human beings have some cognition of the world, where cognition by itself—mere cognition—is weaker than knowledge, the argument proceeds to show that the conditions that have to be satisfied for mere cognition to take place are such as to ensure the possibility of bona fide knowledge, including knowledge of causal relations and laws of nature.

Thus understood, the *Critique* establishes the possibility of *human knowledge* by analyzing the *structure of human cognition* in light of what we may call “the basic human cognitive situation”.¹⁵ The main idea is that for human beings to obtain *any* cognition of the world, certain conditions on the *structure* of human cognition must be satisfied, and these conditions are such as to make not just mere cognition, but also *full-fledged knowledge* (including knowledge of laws of nature), possible.

This is my proposed understanding of the overall structure of Kant's project in the *Critique*. How does truth enter into this project? Truth is one of the central conditions that distinguish knowledge from cognition. Mere cognition is not required to be true, but genuine knowledge is, or so I will argue.¹⁶

2. Truth in Kant's Project

The view that truth plays a significant role in Kant's epistemic project is supported both by thematic considerations and by (direct as well as indirect) textual evidence.

(a) *Thematic considerations*. Kant's epistemic project is an ambitious project. Kant does not seek to establish scientific knowledge in a weak sense. He is not interested only, or even primarily, in the aesthetic virtues of scientific theories, their practical success, convenience, or simplicity, their mere consistency or coherence. Rather, he is concerned with their *veridicality*. His goal is to establish the possibility/viability of a *true* theory of nature, a theory whose judgments—including causal judgments and scientific laws—are *objectively true*. It is only by establishing the possibility of *genuinely true* judgments and theories that

¹⁵ This explains why, in spite of Kant's concern with the possibility of *knowledge*, his discussion is largely devoted to *cognition*.

¹⁶ This view is shared by, e.g., Rosenkoetter (2009).

Kant's goal can be achieved. This strongly suggests that *truth* is absolutely central to Kant's project.¹⁷

(b) *Textual evidence.* The centrality of truth to Kant's project is supported by textual evidence both from the *Critique* and from the *Prolegomena*.¹⁸

- (1) Quite early in the *Critique*, Kant offers a characterization of formal logic. Its task, Kant says, is to set certain general conditions of *truth*, specifically certain general negative conditions: “[Logic concerns] cognition in respect of its mere form (leaving aside all content)... [It] expounds the universal and necessary rules of the understanding... Whatever contradicts these rules is false... **[T]he agreement of cognition with the general and formal laws of [logic].**... is a *conditio sine qua non*, and is therefore the **negative condition of all truth**” (A59-60/B83-4).¹⁹
- (2) Proceeding to transcendental logic and its two parts, transcendental analytic and transcendental dialectic, Kant characterizes the former—arguably, the most central part of the *Critique*—as “a logic of **truth**” (A62/B87) or “**the logic of truth**” (A131/B170), saying, e.g., that “no cognition can contradict it without at once losing... all **truth**” (A62-3/B87).
- (3) In the Dialectic, too, Kant expresses his interest in *truth*. For example, he argues that the possibility of experience is a *necessary condition* for *truth*: “It is possible experience alone that can give our concepts reality; without it, every concept is only an idea, without **truth** and reference to an object”. (A489/B517)
- (4) Both in the Analytic and the Dialectic Kant speaks about “**empirical truth**” (e.g., A191/B236, A492/B520), saying, for example, that “the relation of cause to effect... is the condition of... [the] **truth**” of “empirical judgments” (A202/B247). Spelled out in more detail, his claim is that “the relation of cause to effect... is the condition of the **objective validity** of our empirical judgments with regard to the series of perceptions, thus of their empirical **truth**” (*ibid.*).

¹⁷ I will explain why Kant cannot make do with, say, justification (*sans truth*) in Section 3.3.

¹⁸ I do not include references to truth in Kant's teaching notes, *Logic*, since the relation of these notes to Kant's theory in the *Critique*, which is the basis for my positive lesson, is not necessarily straightforward. The possibility of a gap is further increased by the fact that these notes were compiled by Kant's students rather than by Kant himself.

¹⁹ Kant also uses the expression “criterion” with respect to logic. But logic, he says, provides us only with a negative, formal criterion of truth, and the idea of such a criterion is unproblematic, in contrast to the idea of a positive, material criterion which is involved in the criterial conception of the theory of truth discussed in Part I.

- (5) Kant also talks about “**transcendental truth**”, saying that it “precedes all **empirical truth** and makes it possible” (A146/B185).
- (6) Throughout the *Critique*, Kant is worried about lack of *truth*. In the first, positive part he is concerned with avoiding “**illusion**” (B 69-71, A63-4/B88, B168), “**mere... imagination (as in dreams and delusion)**” (B278), “**mere... seem[ings]**” (B69), “**fictitiously ascrib[ing]**” things to nature (Bxiv), “**random groping**” (Bxv) as opposed to *true* knowledge, and so on. He calls the second, negative part of the *Critique* the “*logic of illusion*” (A293/B249), and he devotes much of it to warnings about using reason in ways that lead to *non-truth*.
- (7) Positively, Kant is continually concerned with epistemic goals that involve or are closely related to truth. These include “**objectiv[ity]**” (A201/B246, A262/B317), “**objective reality**” (A109, A155/B194, A279/B335), and “**objective validity**” (A 125, B142, A202/B247, A211/B256).²⁰ At least in one place (A125) Kant explicitly identifies “**objective validity**” with “**truth**”.²¹
- (8) Kant appeals to *truth* in explaining both his table of judgments (see A74-6/B99-101) and his table of categories. For example, he says that it is implicit in the latter that “[i]n all cognition of an object... there is **truth**, in respect of its consequences” (B114) and that “the **truth** of all that may be immediately deduced from” a concept (along with its “*unity*” and “*completeness*”) is “required for” its full “construction” (B115).
- (9) In the *Prolegomena*, Kant views the Humean challenge that partly motivates, or serves as a catalyst for, his own project as a challenge involving *truth*: “The question was not whether the concept of cause was... useful...; but whether that concept... possessed an inner **truth** [where “inner” is related to the *necessity* of causal truths]” (Kant 1783: 6-7). He views Hume as concerned with the adequacy of our grounds for “accepting something [especially causal claims] as **true**” (*ibid.*: 25). He argues that if “we assume that [scientific] cognitions from pure reason actually exist”, such cognitions “can show the **truth**” of necessary scientific judgments. (*ibid.*: 26). He further says that “[w]hen an appearance is given us, we are still quite free as to how we should judge the matter. The appearance depends upon the senses, but the judgment upon the

²⁰ I limited my citations of the above expressions to cases where they are used with respect to cognitions and judgments (i.e., truth bearers) and their use is positive. Very often they are used with respect to concepts, or negatively. In those cases they are usually related to truth as well, albeit indirectly.

²¹ Hanna (2000: 231, 2004/13) says that Kant views objective validity only as a necessary condition of truth. But other interpreters regard Kant as *identifying* objective validity with truth (See, e.g., Paton 1936, Vol. II, p. 92, fn. 7.).

understanding; and **the only question is whether in the determination of the object there is truth or not**" (*ibid.*: 38). And in the Appendix he emphasizes that his own theory, unlike Berkeley's, provides a "certain criterion for **distinguishing truth from illusion**" (*ibid.*: 124).²²

All these thematic and textual considerations support the view that truth is central to Kant's epistemic project.

But which conception of truth is central to Kant's project: the correspondence conception? the coherence conception? some other conception? In answering this question I will confine myself to truth in the scientific domain. Using Kantian terminology, I will limit my attention to synthetic truth, both a posteriori and a priori, in that domain. The reason for focusing on scientific truth is that the positive lesson I wish to draw from Kant is based on his treatment of this kind of truth. I believe that this lesson can be extended to truth in all domains, but I will not discuss this in the present paper.²³

3. *Truth as Correspondence*

What is Kant's conception of truth? The view that Kant is a proponent of the coherence conception of truth is quite fashionable today, but I side with those who (like Paton 1936; Van Cleve 1999; and Rosenkoetter 2009) view him as a proponent of the correspondence conception. Before presenting my reasons for this view, let me briefly clarify what I do (and do not) mean by "correspondence" and "coherence" in this paper.

Correspondence: By "the" or "a" "correspondence conception of truth" I do not mean the traditional view that a true cognition or judgment is a "copy" or a "picture" or a "mirror" of reality, or even that it stands in the simple relation of "isomorphism" to the facet of reality it targets.²⁴ This view is too simplistic for Kant and, in my view, too simplistic period. The core idea of correspondence, as it applies to Kant—and as, I believe, it ought to apply to contemporary theories of truth as well—is that to be true is to be correct about the *world* and that a true judgment must *be connected to relevant aspects of the world in an appropriate way*. Thus viewed, both the structure of the correspondence relation (appropriate connection) and the aspects of the world that are relevant for truth in different fields are left open. That is, they are treated as open questions, something to be investigated rather than

²² Here, too, Kant appears to be using "criterion" in the general intuitive sense of this word, which differs from his more specific use in the passage cited in Part I.

²³ For an extension of my positive lesson from Kant to mathematical truth, see Sher (2016a, Chapter 8, Section 4).

²⁴ In this paper I use "reality" as synonymous to "world".

determined in advance. And Kant, I will argue, has some important things to teach us about these open questions.

Coherence: By “the” or “a” “coherence conception of truth” I mean the core idea of most coherence theories, namely, that to be true is to *agree* or *cohere* with some body of *judgments* (cognitions, beliefs), where this is usually taken to be the body of our own judgments.

The main philosophical difference between the coherence and the correspondence conceptions of truth is that the coherence conception focuses on relations between *judgments* while the correspondence conception focuses on relations between judgments and relevant aspects of the *world*. Of course, agreement between judgments could involve the world, and likewise, agreement with the world could result in agreement between judgments, but the *center of gravity* of coherence is significantly different from that of correspondence.

3.1. Correspondence

Kant’s “official” definition, or characterization, of truth, which he “assume[s] as granted”, is a *correspondence* characterization: “**truth... is the agreement of cognition with its object**” (A58/B82). This characterization he repeats time and again: “**truth, that is, agreement with the object**” (A157/B197), “**truth consists in the agreement of cognition with its object**” (A191/B236), “**truth (that is... the agreement of our cognition with objects)**” (A237/B296), “**Truth and error... are only to be found... in the relation of the object to our understanding**” (A293/B350), “**truth, that is,... the conformity of our concepts with the object**” (A642/B670), “**truth [is] conformity of the cognition to the object in concreto**” (Kant 1783: 26), and so on.

Moreover, the overall nature of Kant’s ambitious epistemic project strongly suggests a correspondence conception of truth. Kant seeks to establish the possibility of human knowledge in a strong sense of “knowledge”, i.e., genuine knowledge of the *world*. The standard of truth he needs in order to bring this project to fruition is, thus, a strong truth standard, one that connects true cognitions/judgments to the world in a significant and systematic manner. Without a strong normative investment in a standard geared toward the world, Kant’s ambition to do what (he believes) his predecessors failed to do—establish the possibility of genuine knowledge of the world, including scientific knowledge, i.e., knowledge of causal relations and laws of nature—would be seriously compromised.²⁵

The view that Kant’s conception of truth is a correspondence conception is further supported by Kant’s specific choice of expressions to describe his project and argue for his theory. Kant repeatedly invokes many of the cornerstones of the correspondence conception: *reality, objects, representation, existence, givenness, reference, corresponding,*

²⁵ In section 3.3. below we will see why this standard cannot be an epistemic standard in the narrow sense delineated above.

objective, and so on. There are too many occurrences of these notions in the *Critique* to list them here. The number of uses of “reality”, “object”, and “representation” is especially vast. To mention just a few examples:

“*Reality*”. We have already seen that Kant often talks about reality as *objective*. Sometimes he explicitly identifies “**objective reality**” with “**relation to an object**” (see, e.g., A109, A155/B194). Sometimes he merely connects the two in some significant way (see, e.g., A197/B242). He says that the “*objective reality*” of the categories requires “*outer intuition*” (B291). And so on. The centrality of *reality* for his theory is also partly reflected in the fact that it is one of his categories (A80/B106).

“*Object*” and “*representation*”. Kant often talks about *objects* and *representations* in ways that suggest the mind–world relation characteristic of correspondence:

“[H]ow should our faculty of cognition be awakened into action did not **objects affecting our senses**... produce **representations**” (B1), “a mode of cognition... **relate[s] to objects**” (A19/B33), “we are **affected by objects**” (*ibid.*), “an **object**” has “**effect... upon the faculty of representation**” (A19/B34), a “**representation**... when it is an intuition, is in immediate **relation to an object**” (A68/B93), “[a] concept is **related to** [some ... representation of] **an object**” (*ibid.*), “a... **representation**” can “**immediately relate... to an object**” (*ibid.*), “our mind... can **receive representations of objects**” (A77/B102), “the understanding... [is] a faculty of *cognition* that is meant to **relate to objects**” (A97), “**objects... present themselves to our senses**” (B159), “**representations** are... **referred** by the understanding to... **object[s]**” (A250), etc.

“*Existence*”, “*givenness*”, and “*correspondence*”. Kant talks about *existence* and *givenness* in ways that fit in with, or suggest, correspondence, and sometimes he talks directly about correspondence (in a sense pertaining to “correspondence theory of truth”). For example, he talks about “**existence of things outside me**” (B276), “**objects**” being “**given to us**” and “thought” by us (A15/B29), for a “cognition... to have **objective reality**... the **object** must be able to be **given** in some way” (A155/B194), “**objects corresponding to... concepts**” (A224/B272), and so on. “Existence”, like “reality”, is also one of Kant’s categories (A80/B106).

One potential obstacle to viewing Kant as a correspondence theorist is his idealism. As a self-described idealist, how can Kant hold a correspondence notion of truth? However, Kant’s idealism is of a special kind, and he steadfastly emphasizes its difference from “ordinary” (Kant 1783: 124) or “*material*” (B274) idealism—the kind of idealism that threatens the correspondence conception of truth. Material idealism “**declares things in space to be merely imaginary**” (*ibid.*); but Kant

adamantly rejects material idealism. This idealism, which he attributes to Berkeley, “consists in the assertion that there are none but thinking beings, all other things which we think are perceived... being nothing but representations in the thinking being, to which **no object external to them in fact corresponds**” (Kant 1783: 36). Contrasting his own doctrine with this idealism, Kant says:

I, on the contrary, say that things as objects of our senses existing outside us are given... I grant by all means that there are bodies without us, that is, things which... we... know by the representations which their influence on our sensibility procures us. [*Ibid.*]

In the *Critique* Kant devotes two subsections, **Refutation of Idealism** and the Fourth paralogism, to an argument against material idealism. Indeed, so serious is he about it that he offers a “[p]roof” of “*the existence of objects in space outside me*” (B275). Kant describes himself as an empirical realist: “[E]mpirical realism is beyond doubt, i.e., to our outer intuitions there corresponds something real in space” (A375). Similarly, he affirms “the *empirical reality* of time” which, he claims, has “**objective validity in regard to all objects that may ever be given to our senses**” (A35/B52). In an important sense, Kant’s own type of idealism, *transcendental idealism*, is a *means* for achieving empirical realism. Kant is very clear on the relation between transcendental idealism and empirical realism, declaring: “**The transcendental idealist is... an empirical realist**” (A371).

In the *Prolegomena* Kant returns to the question of idealism and realism again and again. “Can this be termed idealism?” (Kant 1783: 36), he asks about his so-called transcendental idealism. His answer leaves no room for doubt: “**It is the very contrary**” (*ibid.*). And again: “my principles... are... **far from turning the truth of experience into mere illusion**” (*ibid.*: 40); “[m]y protestation... **against all charges of idealism is... valid and clear**” (*ibid.*); “I have... given... my theory the name of transcendental idealism, but that cannot authorize anyone to confound it... with... empirical idealism” (*ibid.*); “[m]y **idealism concerns not the existence of things** (the doubting of which... constitutes idealism in the ordinary sense)” (*ibid.*: 41), “it never came into my head to doubt [the existence of things]” (*ibid.*). Indeed, transcendental idealism, according to Kant, undermines empirical idealism and is indispensable for *strong* realism: “**My so-called** (properly critical) **idealism** is of quite a special character, in that it **subverts the ordinary idealism**” (*ibid.*: 124). And not only does it subvert ordinary idealism with respect to empirical knowledge, it also subverts it with respect to a priori knowledge: it is “only through” transcendental idealism that “all *a priori* knowledge... receives **objective reality**” (*ibid.*).

To understand how transcendental idealism subverts ordinary idealism and supports correspondence, it would be helpful to turn to Kant’s Copernican revolution. On the one hand, Kant’s revolution

puts the structure of *human cognition* at the center of his project. But on the other hand, the project itself seeks to establish the possibility of *genuine knowledge of the world*. The key to understanding the centrality of cognition for Kant's project is the fact that cognition is cognition of the *world*. As such, cognition is not possible without the *existence* of both *the world* and *human faculties for cognizing it*. Understanding human cognition involves understanding its connection with the world, and it is this connection that is the basis for the possibility of genuine knowledge. To arrive at correspondence, we note that (for Kant, as for us) not just any cognitive connection to the world is knowledge. To provide genuine knowledge, cognition must satisfy certain constraints, and truth is one of these constraints. The truth constraint requires an *appropriate* connection between genuine knowledge (successful cognition) and the world, and correspondence is just such a connection. But not all conceptions of correspondence would fit Kant's Copernican revolution. As we have noted above, the simplistic conception of correspondence as exhibiting the pattern of *copy* or even *isomorphism* will not. In section 6 I will discuss some of the characteristic features of Kantian correspondence, but first let us consider other views of Kant's conception of truth.

3.2. Coherence

The fact that traditional correspondence—correspondence as “copy”, “mirror”, or “isomorphism”—does not fit into Kant's account of knowledge has naturally led some philosophers to conclude that Kant's theory of truth is not a correspondence theory at all. This further led them to look for other conceptions of truth to explain his view. The most common candidate for such a conception has been the coherence conception of truth. But the coherence conception has other central features besides *not being* a copy, mirror, or isomorphism conception. In particular, the coherence conception replaces the emphasis, central to correspondence, on connections between judgments/cognitions and the world with emphasis on connections between judgments/cognitions and (other) judgments/cognition, ignoring, or at least devaluing, the role played by the *world* in truth.

Is the view that Kant's conception of truth is coherentist supported by solid evidence? I believe it is not. To avoid lengthy polemics, I will limit myself to two points. First, I will show that what appears to be the strongest textual evidence for the view that Kant's conception of truth is coherentist in fact provides no evidence for this view. Second, I will show that the most influential interpreter who claims that Kant is a coherence theorist, Kemp Smith, offers no evidence for his claim.

Starting with textual evidence, there is one paragraph in the *Critique*, in the subsection Opining, Knowing, and Believing of The Canon of Pure Reason, that purportedly suggests a coherence approach to truth:

Taking something to be true is an occurrence in our understanding that may rest on objective grounds, but that also requires subjective causes in the mind of him who judges. If it is valid for everyone merely as long as he has reason, then its ground is objectively sufficient, and in that case taking something to be true is called *conviction*. If it has its ground only in the particular constitution of the subject, then it is called *persuasion*.

Persuasion is a mere semblance, since the ground of the judgment, which lies solely in the subject, is held to be objective. Hence such a judgment also has only private validity, and this taking something to be true cannot be communicated. Truth, however, rests upon agreement with the object, with regard to which, consequently, the judgments of every understanding must agree... The touchstone of whether taking something to be true is conviction or mere persuasion is therefore, externally, the possibility of communicating it and finding it to be valid for the reason of every human being to take it to be true; for in that case there is at least a presumption that the ground of the agreement of all judgments, regardless of the difference among the subjects, rests on the common ground, namely the object, with which they therefore all agree and through which the truth of the judgment is proved. [A820/B848]

Now, if we remove from this paragraph everything that suggests a correspondence reading, we obtain a paragraph that may naturally be read as expressing a coherence view of truth. (For emphasis I use italics for those parts that appear to suggest coherence:)

Taking something to be true is an occurrence in our understanding that requires subjective causes in the mind of him who judges. If it is valid for everyone[,] then its ground is objectively sufficient, and in that case taking something to be true is called conviction. If it has its ground only in the particular constitution of the subject, then it is called persuasion.

Persuasion is a mere semblance, since the ground of the judgment, which lies solely in the subject, is held to be objective. Hence such a judgment also has only private validity[.]

[W]ith regard to [truth],
the judgments of every understanding must agree. . .

The touchstone of whether taking something to be true is conviction or mere persuasion is therefore the possibility of communicating it and finding it to be valid for the reason of every human being to take it to be true; in that case there is agreement of all judgments, regardless of the difference among the subjects[.]

[T]herefore
[T]he truth of the judgment is proved.

[A820/B848]

But if we shift our attention to the parts that express the *basic ideas* of the original paragraph, we immediately see that Kant's focus is reversed. Using boldface for ideas that, in the context of this paragraph, are correspondence ideas, italics for coherence ideas, underlining for the relations between the two, small caps for parts of the paragraph which were emphasized in the original, and unmarked typeface for the rest, will clearly show that for Kant coherence is relevant *only in being a sign* for satisfaction of the *real ground* of truth—*correspondence*.

Taking something to be true is an occurrence in our understanding that may rest on objective grounds, but that also requires subjective causes in the mind of him who judges. If it is valid for everyone merely as long as he has reason, then its ground is objectively sufficient, and in that case taking something to be true is called CONVICTION. If it has its ground only in the particular constitution of the subject, then it is called PERSUASION.

Persuasion is a mere semblance, since the ground of the judgment, which lies solely in the subject, is held to be objective. Hence such a judgment also has only private validity, and this taking something to be true cannot be communicated. **Truth, however, rests upon agreement with the object, with regard to which, consequently, the judgments of every understanding must agree. . . The touchstone of whether taking something to be true is conviction or mere persuasion is therefore, externally, the possibility of communicating it and finding it to be valid for the reason of every human being to take it to be true; for in that case there is at least a presumption that the ground of the agreement of all judgments, regardless of the difference among the subjects, rests on the common ground, namely the object, with which they therefore all agree and through which the truth of the judgment is proved.**

[A820/B848]

As we can see from the new highlighting, the conception of truth expressed in this passage is a *correspondence* conception, coherence serving as no more than an “**external**” sign that *correspondence* is satisfied. Given that throughout most of the *Critique* Kant expresses a correspondence conception of truth without reference to coherence at all, even this secondary role of coherence is marginalized.

Turning, next, to exegesis, the view that Kant’s conception of truth is a coherence conception is often traced to Kemp Smith’s *Commentary to Kant’s “Critique of Pure Reason”* (1918/23/84). It is important to realize, however, that Kemp Smith says very little about this issue in the *Commentary* and that his claims are not supported by any textual or thematic evidence. The only text Kemp Smith appeals to as allegedly expressing a commitment by Kant to the coherence conception of truth is the Introduction to the *Critique*. In his commentary on the Introduction, Kemp Smith says:

If our account of Kant’s awakening from his dogmatic slumber be correct, it consisted in his recognition that self-evidence will not suffice to guarantee any general principle. The fundamental principles of our experience are synthetic. That is to say, their opposite is in all cases conceivable. Combining this conclusion with his previous conviction that they can never be proved by induction from observed facts, he was faced with the task of establishing rationalism upon a new and altogether novel basis. If neither empirical facts nor intuitive self-evidence may be appealed to, in what manner can proof proceed? And how can we make even a beginning of demonstration, if our very principles have themselves to be established? Principles are never self-evident, and yet principles are indispensable. Such was Kant’s unwavering conviction as regards the fundamental postulates... of knowledge. [*ibid.*: 36]

So far, there is no mention of, or connection to, coherence. Then, however, out of the blue, Kemp Smith declares:

This is only another way of stating that Kant is the real founder of the *Coherence* theory of truth. [*ibid.*]

In the absence of *any* explanation of *how* the paragraph cited above presents Kant as a “founder” of the coherence theory of truth, this declaration is a *non-sequitur*.

This conclusion is strengthened by Kemp Smith’s admission that Kant himself “**never... employs the term Coherence**” and, moreover, that “**he constantly adopts positions which are more in harmony with a *Correspondence* view of the nature and conditions of knowledge**” (*ibid.*). Still, there must be some explanation for Kemp Smith’s declaration. The clue is given in the next sentence. Following his admission that Kant himself adopts correspondence views, Kemp Smith

says: “**But** all that is **most vital** in his teaching, and has proved really fruitful in its **after-history**, would seem to be in line with the **positions which have since been more explicitly developed** by such writers as Lotze, Sigwart, Green, Bradley, Bosanquet, Jones, and Dewey, and which **in their tenets all derive from Hegel’s restatement of Kant’s logical doctrines**” (*ibid.*). Since these *later* positions are “**now usually entitled the Coherence theory of truth**” (*ibid.*: xxxvii), it is *justified* to claim that Kant *is* a coherence theorist.

This kind of justification, however, has little force. It is one thing to say that later philosophers who were influenced by Hegel’s “restatement” of Kant’s view endorsed a coherence theory of truth, and another to declare, as coherentist interpreters of Kant do, that Kant himself is a coherence theorist.

Another place where Kemp Smith talks about Kant and the coherence theory of truth is in the Introduction to the *Commentary*. There he says that “[t]he fundamental thesis of the Coherence theory finds explicit formulation in Kant’s doctrine of the judgment” (*ibid.*: xxxviii).

This theory Kemp Smith describes as “the doctrine, that awareness is identical with the act of judging, . . . [and that] [n]ot contents alone, but contents interpreted in terms of some specific setting, are the sole possible objects of human thought” (*ibid.*). But it is far from clear that awareness or interpretation is incompatible with correspondence, and Kemp Smith offers no explanation, let alone argument, in support of his view.

3.3. *Narrowly Epistemic Conception*

The possibility that Kant’s conception of truth is a narrowly epistemic conception has also been suggested in the literature:

Although Kant does not put it this way, . . . we can view him as rejecting the idea of truth as correspondence . . . and as saying that the only sort of truth we can have an idea of, or use for, is *assertibility* (by creatures with our rational natures) *under optimal conditions* (as determined by our sensible natures). **Truth becomes a radically epistemic notion.** [Putnam 1983: 210]

If we take assertibility under optimal conditions to be, at bottom, justification, then the view is that truth is reducible to justification. But the view that truth is reducible to justification is generally problematic, and it is especially problematic for Kant. The general problem is that justification requires a *standard*. Take the justification of vegetarianism as an example. One can justify one’s vegetarianism by reference to a moral standard, a pleasure standard (taste), a well-being (e.g., health) standard, etc., but one cannot justify it without appeal to any standard.

Similarly, one can only justify one's acceptance of a scientific theory by reference to some standard: a standard of correspondence to the world, a standard of practical utility, etc. So to say that truth is justification is not to say that truth is a matter of practical utility *rather than* of correspondence with the world. And to say that truth is reducible to justification simpliciter is empty. Now, an empty, or a relatively weak, notion of justification is especially problematic for Kant, since his ambitious project requires a substantial and relatively strong type of justification, indeed one that is guided by the goal of a *correct depiction of the world* (rather than practical utility). This consideration, together with the absence of textual evidence for a narrowly epistemic conception of truth in the *Critique*, suggests that Kant is *not* a narrowly epistemic theorist of truth.

Two preliminary issues I would like to address before turning to the correspondence theory of truth incipient in the *Critique* are: (i) Kant's *negative* versus *positive* orientation, and (ii) Kant's *transcendental* methodology.

4. *Positive Orientation of the Critique*

My goal is to derive a *positive* lesson from Kant on truth based on the *positive* content of the *Critique*. But Kant often describes the *Critique* as performing a *negative* role, namely, that of identifying the *limits* of human knowledge. Does the *Critique* have sufficient positive content to support a positive lesson?

I believe it does. As I understand it, the negative orientation of the *Critique* is subservient to its positive orientation. The underlying idea is that in order to establish the possibility of human knowledge, we have to recognize the scope and limits of such knowledge. If we exaggerate its scope, it would be impossible to establish its possibility, simply because it is impossible to establish what is not possible. The negative task of the *Critique* is, thus, to identify the limits of human knowledge, and the history of philosophy taught Kant that the crucial point is the limitation of *pure reason*. This is reflected in the title of his work: "Critique of Pure Reason". In humans, reason, by itself, is incapable of providing genuine knowledge of the world, and it is only by taking this fact into account that we are able to establish the possibility of human knowledge. The negative aspect of the *Critique* is thus indispensable for its positive goal.²⁶

²⁶ Kant calls the kind of knowledge whose possibility he seeks to establish "theoretical knowledge". Pure reason, according to him, is capable of providing humans with *practical* knowledge as well as with guidelines for theoretical knowledge ("regulative principles"), but not with actual theoretical knowledge.

Linguistically, the relation between the negative and positive sides of the *Critique* is reflected in the use often made by Kant of expressions like “only”. For example, when Kant makes the *negative claim* that “time... has objective validity **only** in respect of appearances” (A34/B51), he also makes the *positive claim* that time *has* objective validity in respect of appearances. When he says that “[o]nly through the... union [of “the understanding” and “the senses”] can cognition arise” (A51/B75-6), he also says that through the union of the understanding and the senses cognition *can* arise. When he says that “receptivity can make cognition possible **only when** combined with spontaneity” (A97), he also says that *when* receptivity is combined with spontaneity it *can* make cognition possible; and so on.

In both direct and indirect ways, therefore, the *Critique*, particularly in the Aesthetic and the Analytic, includes a positive theory of the structure of human cognition and its relation to the world. This theory, Kant believes, demonstrates the possibility of human (theoretical) knowledge, and, on the present interpretation, it includes an implicit positive theory of truth.

5. Kant's Transcendental Methodology

One distinctive feature of Kant's theory of knowledge is its methodology, and in particular the standpoint from which it approaches its subject matter, i.e., the level at which theorizing is conducted. Kant's epistemic theory studies its subject matter from a *transcendental viewpoint*, i.e., Kant's theorizing is conducted on a *transcendental level*. This is directly relevant to our inquiry, since if Kant's theory of knowledge embeds, either explicitly or implicitly, a substantive theory of truth, or at least the seeds of such a theory, this theory is likely to be *transcendental* as well. And if this theory is a *correspondence* theory, then it is likely to be a *transcendental correspondence theory*.

To draw a positive lesson from Kant on the viability of a substantive correspondence theory, we need to understand what being *transcendental* amounts to. This requires that we heed Kant's warning about confusing “*transcendental*” with “*transcendent*” (see, e.g., A296/B352-3 and B427). To this end I will introduce a new pair of terms: “HG-transcendence” and “HH-transcendence”. Briefly, we can explain the meaning of these terms as follows:

HG-Transcendence: Transcending a *human* standpoint, X, to an ultimate, *Godly*, standpoint Y.

HH-Transcendence: Transcending one *human* standpoint, X, to another, *human*, standpoint Y.

Using the second of these notions, we can characterize a transcendental theory of X (in Kant's sense) as follows:

Transcendental Theory:

1 A transcendental theory is *HH Transcendent*.

2 A transcendental theory of X focuses on the conditions of the possibility of X.

In contrast, a *transcendent* theory (in Kant's sense) is HG transcendent, and its goal is to provide knowledge of the world as it appears from a God's eye view.²⁷ Thus, while a *transcendental* theory seeks to establish the possibility of human knowledge within the bounds available to humans, a *transcendent* theory, which does not recognize these limits, seeks to provide knowledge of the world from a Godly viewpoint. But such knowledge is not available to humans, Kant insists. Contemporary philosophers have made a similar point.²⁸

This suggests that the theory of truth incipient in the *Critique* approaches truth from an HH-transcendent standpoint, that is, from a *fairly high* standpoint, though importantly from a *human* standpoint. From this standpoint it examines the *conditions under which human judgments/cognitions in general, and scientific judgments/theories in particular, can be true*, and this perspective affects its conception of the correspondence relation underlying truth in a way that will be discussed below.

6. Kant's Incipient Theory of Truth

6.1. The Question of Truth

One of the central questions of the *Critique* is whether *nature* can be *known* by us. And implicit in this question is the question whether *scientific theories*, and in particular scientific *laws* (or statements of such laws), can be *true*. The problem, as Hume pointed out, is that such laws are both necessary and universal, but neither necessity nor universality can be based on actual sensory experience. Expressed in terms of truth, the problem is that while correspondence with particular observable facts is sufficient for particular-and-contingent truth, it is not sufficient for general-and-necessary truth. Kant's solution to the problem is to develop a new cognitive account of knowledge of

²⁷ See, e.g., A690-1/B718-9 and A845-6/B873-4.

²⁸ For Kant's warnings against seeking a transcendent standpoint see A297/B354, A562/B590, A643/B671, A702/B730, as well as Kant 1783: 76 and 99. For a contemporary philosopher's warning against seeking such a standpoint, see Putnam (1981, 1983).

scientific laws. Such knowledge, he says, is not a posteriori, as Hume assumed it was, but *a priori*. More precisely, judgments of scientific laws are, in his terminology, “synthetic a priori”. This explains why he identifies the basic question of the *Critique* as: “**How are synthetic a priori judgments possible?**” (B73). This question is, at bottom, both a question of truth and a question of knowledge: How are *true* synthetic a priori judgments possible? How is it possible for *true* synthetic a priori judgments to be *known*? The key to answering this question, for Kant, is his new, transcendental account of human cognition, which includes a new account of the relation between mind and world. Let us turn to this account.

6.2. *The Structure of Human Cognition*

The two basic components of human cognition, according to Kant, are “receptivity” and “spontaneity” (A50/B74), and its two basic sources are *mind* and *world*. In a simple act of cognition, the mind *receives* stimuli from some source external to it (receptivity), and it *shapes* these stimuli using its own internal resources (spontaneity). The external element is *world* or *reality*; the internal element—*mind*. From a transcendental perspective, each of these elements is itself complex. Reality has two levels: (i) absolute reality—“thing in itself” (Bxx) or “noumenon” (B307), and (ii) experiential reality—“appearance” (Bxx) or “phenomenon” (A183/B227). The mind has three faculties: (i) *sensibility or intuition* (which itself has two components: sensory intuition and pure intuition), (ii) *imagination*, and (iii) *understanding*.

Roughly, and without aiming at a comprehensive description, we may sketch the stages involved in generating scientific cognitions as follows²⁹: First, external input is received through our sensory apparatus. Next, this input is shaped by our (sensory and pure) intuition. The results then undergo a chain of syntheses, starting with syntheses generated by our imagination and continuing with syntheses produced by our faculty of concepts—understanding. In each stage we achieve some level of *representation of the world*, and representations generated in lower stages can be further synthesized in higher stages. Representations of full-fledged *objects* are achieved only at relatively late stages. That is, it is only after our sensory input is shaped by our forms of intuition and undergoes synthesis by imagination and concepts that full-fledged *objects* emerge. (Full-fledged) objects, thus, reside in a level of reality which is partly shaped by us; but they are anchored in another, deeper, level of reality which is completely independent of us. The former level of reality is the one Kant calls

²⁹ Given the multiplicity and diversity of interpretations of Kant, the present sketch cannot agree with all the existent interpretations. Nor will it fully coincide with any one of these. Still, many of these interpretations share sufficient features with our sketch that our positive lesson from Kant will be pertinent to them.

“appearance” (“phenomenon”, “experiential reality”), the latter —“thing in itself” (“noumenon”, “the *unconditioned*” (Bxx), “absolute reality” (A35/B52)). It is a central tenet of Kant’s theory that “[w]hat the objects may be in themselves would never become known to us” (A43/B60) and that the level of reality in which they are accessible to us is the level of appearances. It is this level of reality that scientists refer to as “Nature” and that science provides us with knowledge of. “[N]ature”, in Kant’s words, is “the sum of all appearances” (B163). But the level of thing in itself is also indispensable, since “otherwise we should be landed in the absurd conclusion that there can be appearance without anything that appears” (Bxxvi-vii).

Our cognition of nature takes many forms: from cognition of individual objects and their contingent properties (relations) to highly general and abstract cognitions, cognitions of universal and necessary laws of nature on various levels. An example of cognition of the former kind is a cognition of a certain book being heavy or of one book lying next to another. An example of a cognition of the latter kind is cognition of general-and-necessary laws like “All alterations take place in conformity with the law of the connection of cause and effect” (B232).

6.3. *The Theory of Truth*

Our discussion so far suggests that the theory of truth embedded in the *Critique* is a transcendental correspondence theory. Its goal is *neither* to provide a mere “name-clarification” (nominal definition) of truth *nor* to provide a “criterion” of truth (in the sense discussed in Part I). And its focus is *neither* on a compositional account of truth, as in Tarski’s or Davidson’s theory, *nor* on the truth predicate as an instrument of generalization, as in deflationist theories.

Taking a relatively broad HH-transcendent perspective on human cognition, Kant’s implicit theory of truth focuses on the basic structure of truth as it connects human judgments—in particular, scientific judgments—to the world. Truth depends both on the way the *world* is and on the way our *mind* operates, and given the complexity of the human cognitive situation, the structure of both world and mind, as they pertain to truth and knowledge, is *complex*.

One reflection of our complex cognitive situation is that what is *appearance* from the transcendental philosopher’s point of view is *reality as it is in itself* from the scientist’s point of view. Thus, while from the scientist’s narrower viewpoint scientific truths correspond to *things as they are*, from the transcendental philosopher’s broader viewpoint they correspond to *things as they appear to beings with our cognitive nature to be*. Take Kant’s example of *rain*. Due to having a narrower perspective, the scientist views “rain [as a] thing in itself” (A45/B63). “This is [a] correct [view]” from her perspective, since as a scientist she takes the concept of thing in itself “in a merely physical sense” (*ibid.*). From this perspective, rain is a fixed thing, since it is always “determined”

in the same way “in our intuition” (*ibid.*), that is, in the same way relative to “all [our] experience[s] and in all its various positions relative to [our] senses” (*ibid.*). Viewing rain from this perspective, therefore, the scientist’s claims about rain, as well as about natural laws applicable to rain, are *objective*. But the transcendental philosopher “take[s] this empirical object [rain] in its general character, and ask[s]... whether it represents an object in itself” (A45-6/B63). Here, “the question as to the relation of the representation to the object at once becomes transcendental”, and from a transcendental perspective we immediately “realise that not only are the drops of rain mere appearances, but that even their round shape, nay even the space in which they fall, are nothing in themselves, but merely modifications or fundamental forms of our sensible intuition, and that the transcendental object [thing in itself] remains unknown to us” (*ibid.*).

Accordingly, Kant’s theory offers a two-dimensional account of truth: an account of the nature of truth when viewed from within science, and an account of the nature of truth when viewed from a relatively high HH-transcendent standpoint, external to science—one that has both science itself and the world it studies in view. Each account throws partial light on truth; together they throw a fairly comprehensive light on it.

From both perspectives, truth is *correspondence* with reality. Viewed from within science, truth is correspondence with reality in a relatively simple way; viewed from a transcendental perspective, truth is correspondence with reality in a rather complex way. From the latter perspective, scientific truths correspond to reality both on the level of appearance and on the level of thing in itself, but (according to Kant) they correspond to appearances in a way that is analyzable by us and they correspond to things in themselves in a way that is unanalyzable by us. Thus viewed, what we may call “the Kantian route of correspondence” has two segments. Its first segment connects human judgments to appearances; its second segment connects appearances to things in themselves. Using contemporary terminology, we may say that a scientific judgment of the form “All A’s are B’s” is true if and only if (i) the extension of “A” on the level of appearance is included in the extension of “B” on the level of appearance, and (ii) the level of appearance is appropriately grounded in the level of thing in itself. The details of Kant’s grounding of appearances in things in themselves are in dispute among commentators, and the very idea of a thing in itself as an ultimate layer of reality is contested by many philosophers. But for the purpose of the present discussion there is no need to decide these issues. What is important from our perspective is that the Kantian “route of correspondence” is complex and that there is a difference in complexity between the route of correspondence as viewed from within science and as viewed from a transcendental, philosophical perspective.

7. A Positive Lesson on the Structure of Truth from Kant

We are finally ready to draw a positive lesson on the structure of truth for contemporary theorists. In drawing this lesson, it is important to keep in mind that it is intended for contemporary philosophers in general, not just those whose philosophical positions are close to Kant's. Accordingly, the lesson has to avoid commitment to many of Kant's distinctive tenets. Sometimes this can be done by offering new analysis of these tenets (see above). But more generally, this can be achieved by treating the *Critique* as a *model*, or an *example*, of how a certain approach to truth can be realized, leaving room for the development of other models of this approach. As a model, Kant's theory establishes, or at least strongly suggests, the possibility of a genuinely substantive theory of truth and, upon appropriate *generalization*, provides insight into the complexity of the correspondence relation. What forms this complexity can and should take is independent of the precise details of Kant's model.

Not surprisingly, our positive lesson from Kant is not a one-line lesson, but a cluster of points that, together, make up a largely positive lesson. Without purporting to provide a complete or an all-inclusive statement of these points, we may briefly and provisionally formulate them as follows, starting with truth and knowledge in science:

Positive Lesson on Truth from Kant:

- (A) Full-fledged scientific knowledge, one that encompasses laws of nature, or significant regularities in nature, requires a *standard of truth or correctness*, in addition to standards of evidence and justification.
- (B) Such a standard sets substantial demands on the relation between *correct cognitions* and relevant aspects of the *world* and as such is a *substantive correspondence* standard.
- (C) As a standard of correctness for *human* cognitions vis-a-vis the world (their target), the standard of truth has to take into account the *human cognitive situation*. This situation is highly *complex*, due to the complexity of *human cognition* on the one hand and the complexity of the *world* (relative to our cognitive capacities) on the other.
- (D) Contrary to common opinion, the dependence of truth on the human cognitive situation does not conflict with genuine (robust, substantive) correspondence. Just because a standard X takes Y into account does not mean that it does not take Z into account as well, possibly to a *considerable* degree. Specifically, just because a standard of truth takes the structure of human cognition into account does not mean that it does not demand a

deep, robust, and systematic connection between true cognitions and the world.

- (E) One of the central tasks of a *substantive* theory of truth is to *investigate the connections between the human mind and the world that are responsible for truth*. This means that the structure of correspondence is unlikely to be as simple as *copy*, *mirror image*, or even *direct isomorphism*. Most importantly, the structure of correspondence in a given field of knowledge cannot be determined in advance, prior to serious investigation of the mind–world relation in the relevant area. Such an investigation need not be purely a priori or transcendental in the precise Kantian sense, but it is expected to be HH transcendental and use intellectual resources that go beyond those afforded to humans according to narrow empiricist conceptions of knowledge (such as Hume’s). Adopting a Wittgensteinian phrase, we may say that the truth theorist should assume the epistemic stance of “look and see”: look and see what the route of correspondence is; investigate, figure out, discover. Furthermore, the focus of a substantive theory of truth is the *central principles* governing truth. This constructive task is vastly different both from the task of formulating a *general schema for generating trivial truth-biconditionals* (deflationist conception of the theory of truth) and from the task of providing a *criterion that determines the truth-value of each and every judgment all by itself* (criterial conception of a substantive theory of truth).
- (F) A substantive correspondence theory of truth does not require HG-transcendence. Accordingly, it is not subject to criticisms of the correspondence theory of truth like those of Putnam (1981, 1983), which focus on its alleged commitment to a “God’s eye view”.
- (G) An adequate substantive correspondence theory of truth requires an attenuated, multi-dimensional conception of the world, but not specifically the Kantian duality of noumena and phenomena. As such it can avoid commitment to Kant’s problematic idea of a thing in itself, contrary to what Putnam (*op. cit.*) and others have said.
- (H) An adequate substantive theory of truth must balance its interest in the generality and diversity/particularity of truth. Accordingly, it will be open to the possibility that different principles of truth vary in their degree of generality.
- (I) Although our positive lesson from Kant is based on his treatment of scientific truth, this lesson may be extended to all types of truth, including logical and mathematical truth, and possibly even moral truth. For an example of an account of logical and

mathematical truth based on a conception of truth in the spirit of (A)-(H) see Sher (2016a).³⁰

This concludes my discussion of two significant lessons on truth from Kant: a largely negative lesson and a significantly positive lesson. One might be tempted to ask: "But what is the exact substantive theory of truth these lessons lead to?" The answer is that these lessons do not determine a unique theory. They leave the details of an adequate substantive theory of truth an open question, but they narrow the space of possibilities and provide constructive clues.

References

- Allison, H.E. 1983. *Kant's Transcendental Idealism*. New Haven: Yale University Press.
- Bennett, J. 1966. *Kant's Analytic*. Cambridge: Cambridge University Press.
- Blackburn, S. 1984. *Spreading the Word*. Oxford: Oxford University Press.
- Cohen, H. 1907. *Kommentar zu Immanuel Kants Kritik der reinen Vernunft*. Leipzig: Durr.
- David, M. 1994. *Correspondence and Disquotatation*. Oxford: Oxford University Press.
- Dyson, F. 1988. *Infinite in All Directions*. New York: Harper & Row.
- Field, H. "The Deflationary Conception of Truth", In *Fact, Science and Morality*, eds. by G. MacDonald and C. Wright. Oxford: Blackwell, 1986 (55–117).
- Field, H. "Deflationist Views of Meaning and Content," *Mind* 103 (1994): 249–285.
- Gardner, S. 1999. *Kant and the Critique of Pure Reason*. London: Routledge.
- Gupta, A. "A Critique of Deflationism," *Philosophical Topics* 21 (1993): 57–81.
- Guyer, P. 1987. *Kant and the Claims of Knowledge*. Cambridge: Cambridge University Press.
- Hanna, R. "Kant, Truth & Human Nature," *British Journal of the History of Philosophy* 8 (2000): 225–250.
- Hanna, R. 2001. *Kant and the Foundations of Analytic Philosophy*. Oxford: Oxford University Press.
- Hanna, R. 2004/13. "Kant's Theory of Judgment." *Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), ed. Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/sum2014/entries/kant-judgment/>.
- Horwich, P. 1990/8. *Truth*. Oxford: Blackwell.
- Hossenfelder, M. 1978. *Kants Konstitutionstheorie und die transzendente Deduktion*. Berlin: De Gruyter.
- Kant, I. 1781/7. *Critique of Pure Reason*. Kemp Smith translation: Palgrave Macmillan (1929–2007). Cambridge: Guyer & Wood translation (1998).
- Kant, I. 1783. *Prolegomena to Any Future Metaphysics*. Beck translation. Macmillan (1950).
- Kant, I. 1770's-1800. *Lectures on Logic*. Cambridge 1992.
- Kemp Smith, N. 1918/23/84. *Commentary to Kant's "Critique of Pure Reason"*. Atlantic Highlands: Humanities.
- Leeds, S. "Theories of Truth and Reference," *Erkenntnis* 13 (1978): 111–129.
- Longueness, B. 1993. *Kant & The Capacity to Judge*. Princeton: 1998.
- Lynch, M. 1998. *Truth in Context*. Cambridge: MIT.
- Lynch, M. 2009. *Truth as One and Many*. Oxford: Oxford University Press.
- Paton, H.J. 1936. *Kant's Metaphysics of Experience*, Vol. I-II. London: George Allen & Unwin.

³⁰ Since (A)-(H) are based on a generalization of Kant's account of scientific truth, the above-noted account of logical and mathematical truth differs from Kant's own accounts of these types of truth.

- Posy, C. 2000. "Immediacy and the Birth of Reference in Kant: The Case for Space." In *Between Logic and Intuition*. Cambridge: 155–185.
- Prauss, G. "The Problem of Truth in Kant," *Kant-Studien* 60 (1969): 166–182 (<http://www.kommentare-zu-gerold-prauss.de/Praussengl.htm#TRUTH>).
- Putnam, H. 1981. *Reason, Truth and History*. Cambridge: Cambridge University Press.
- Putnam, H. 1983. *Realism and Reason*. Cambridge: Cambridge University Press.
- Ramsey, F.P. "Facts and Propositions," *Aristotelian Society Supplementary Volume* 7 (1927): 153–170.
- Rosenkoetter, T. "Truth Criteria & the Very Project of a Transcendental Logic," *Archiv für Geschichte der Philosophie* 91 (2009): 193–236.
- Sher, G. "On the Possibility of a Substantive Theory of Truth," *Synthese* 117 (1999): 133–172.
- Sher, G. "In Search of a Substantive Theory of Truth," *Journal of Philosophy* 101 (2004): 5–36.
- Sher, G. 2016a. *Epistemic Friction: An Essay on Knowledge, Truth, and Logic*. Oxford: Oxford University Press.
- Sher, G. "Substantivism about Truth," *Philosophy Compass* 11 (2016b): 818–828.
- Strawson, P.F. 1966. *The Bounds of Sense*. London: Methuen.
- Tolley, C. Forthcoming. "Idealism and the Question of Truth." *Oxford Handbook on Truth*, ed. by M. Glanzberg. Oxford: Oxford University Press.
- Van Cleve, J. 1999. *Problems from Kant*. Oxford: Oxford University Press.
- Vanzo, A. "Kant on the Nominal Definition of Truth," *Kant-Studien* 101 (2010): 147–166.
- Williams, M. "Meaning and Deflationary Truth," *Journal of Philosophy* 96 (1999): 545–564.
- Wolff, R.P. 1973. *Kant's Theory of Mental Activity*. Gloucester: Peter Smith.
- Wright, C. 1992. *Truth and Objectivity*. Cambridge: Harvard.