Minimalism and the Facts About Truth

Marian David
University of Notre Dame

Minimalism, Paul Horwich’s deflationary conception of truth, has recently received a makeover in form of the second edition of Horwich’s highly stimulating book *Truth*¹. I wish to use this occasion to explore a thesis vital to Minimalism: that the minimal theory of truth provides an adequate explanation of the facts about truth. I will indicate why the thesis is vital to Minimalism. Then I will argue that it can be saved from objections only by tampering with the standards of adequate explanation—a move that deprives it from giving support to Minimalism.

At the heart of Minimalism lies a theory of truth for propositions. It is called the minimal theory, or *MT* for short. It consists of a collection of axioms. Each axiom is a proposition of the form

\[(E) \quad \text{The proposition that } p \text{ is true if and only if } p.\]

MT comprises all propositions of this form, except the ones that give rise to the liar paradox. Note that MT should be distinguished from the schema, (E), used to convey MT, as well as from (E)’s substitution instances, which are sentences rather than propositions. MT consists of all propositions expressed by the sentences that would result from replacing ‘p’ in (E) with a non-pathological declarative sentence of English, or of any possible extension of English—where the non-pathological replacements for ‘p’ are the ones that do not lead to liar-paradoxical substitution instances of (E).²

---

² Cf. Horwich, op. cit., pp. 3-8, 17-22. Horwich thinks it is feasible to restrict the permissible
Minimalism is not confined to MT. It also offers an account of our concept of truth, an account of the meaning of the word ‘true’ (its use and function), as well as minimal theories of utterance-truth and reference. Here I will mostly focus on MT, which is supposed to be a theory of truth itself, of the property of being true. MT is put forward as the correct (albeit implicit) definition of truth. One might object that truth also applies to sentences and utterances. Horwich would respond that the term ‘true’, when applied to sentences or utterances, really means ‘expresses a truth’, i.e., he would maintain that, strictly speaking, truth does not apply to sentences or utterances. Since MT is described as a collection of propositions, Minimalism presupposes the existence of propositions. Those who reject propositions will hold that there is no such theory as MT. Horwich could respond that they are simply wrong about the existence of propositions. Fair enough. However, Minimalism presupposes the existence of propositions in yet another way. According to Minimalism, our grasp of the concept of truth consists in our disposition to accept the substitution instances of (E). But those who reject propositions, as well as those who are merely skeptical about them, will not be disposed to accept all instances of (E). Consider the conditional: ‘If snow is white, then the proposition that snow is white is true’. Its antecedent is true. Since its consequent implies (or presupposes) the existence of propositions but its antecedent does not, the conditional is true only if there are propositions. So, if you reject propositions, or if you are merely skeptical about them, then you will not accept this conditional. Horwich would have to say, implausibly, that proposition-nihilists and proposition-skeptics lack the concept of truth. Better to replace (E) with the following: The proposition that p is true if and

---

3 Cf. Horwich, op. cit., pp. 16-17.
only if the proposition that \( p \) exists and \( p \). Having said this, I will usually stick to Horwich’s original versions for the sake of simplicity.

Horwich claims that the minimal theory, MT, provides an adequate explanation of all the facts involving truth. Let us call this the Adequacy Thesis.\(^5\) This thesis is the main focus of the present paper. Before I discuss it in more detail, I want to explain why I take it to be crucial to Minimalism.

Minimalism holds that MT is the right theory of truth. This contention involves three subclaims: MT is a theory; MT is a theory of truth; MT is the (correct) theory of truth. Each claim faces an initial worry: (a) As theories go, MT seems a bit odd. It does not offer any general principles about truth. Instead, it offers an infinite collection of propositions (the so-called axioms), each one specifying a separate necessary and sufficient condition for the truth of only one particular proposition. Ordinarily, theories are expected to offer more than a bunch of particular propositions; they are expected to convey at least some informative general principles pertinent to their subject matter. So, why does MT deserve to be called a theory at all?\(^6\) (b) MT is a collection of particular truths about truth. But it is equally a collection of particular truths about propositions. So, why is MT a theory of truth rather than a theory of propositions? (c) The axioms of MT seem rather innocuous (assuming one accepts propositions). Surely, they are compatible with almost any theory of truth (for propositions), and most such theories will embrace them gladly. On what grounds, then, can one make the claim that MT is the theory of truth—a claim which implies the strong negative thesis that there are no (correct) theories of truth besides MT? Note especially that the exclusivity of MT is not happily defended

---


\(^6\) MT is an implicit definition of truth, as such it cannot be expected to specify a feature common to all and only those propositions that are true. Still, other theories that are usually regarded as implicit
merely on the grounds of simplicity. While schema (E) is simple enough, MT itself, although rather monotonous, is not very simple. It is, after all, an infinite theory and many of its axioms are enormously complex.\(^7\)

I take it that the Adequacy Thesis is crucial to Minimalism because it figures prominently in the answers to all three questions, giving vital support to the contention that MT is the right theory of truth. Take question (a). A theory should provide us with explanations of all the facts in its domain. It is not too farfetched to consider this service so important that anything providing it deserves to be regarded as a theory. So, if the Adequacy Thesis is correct, MT deserves to be regarded as a theory, even if it is rather odd in other respects. This also answers the first half of (b). With respect to the second half of (b), the minimalist will point out that MT is not a theory of propositions because it does not provide explanations of all the facts about propositions; e.g., it says nothing about the role of propositions in psychological attitudes, like believing or doubting. What about (c)?—probably the most important point. The argument for the exclusivity of MT is largely methodological. It is a best-explanation argument: other theories of truth are to be rejected because they fail to provide (equally) adequate explanations of the facts about truth. Evidently, such a best-explanation argument will not support MT in the absence of the Adequacy Thesis. The explanatory failure of other theories speaks in favor of MT, only if MT does provide adequate explanations of the facts about truth. Below we will encounter a further reason why the Adequacy Thesis is crucial to Minimalism. But first we have to take a closer look at the thesis itself.

---

\(^7\) I have simplified matters a bit. Horwich now claims that “the minimal theory of truth is the theory of truth, to which virtually nothing more should be added”; op. cit., p. 43. What should be added is a single axiom saying that only propositions are true. So Horwich’s complete theory of truth—MT plus this axiom—now does contain one general principle after all. However, the principle is too thin to address the questions raised in the text. (I will suppress reference to this additional axiom when it makes no difference to the discussion.)
Our present formulation of the Adequacy Thesis is potentially misleading: it might be taken to suggest that MT all by itself can explain all the facts involving truth—this would be wildly implausible. But, as Horwich points out, it would also be quite unreasonable to make such a demand on MT. When we say that a theory of X explains the facts about X, we do not expect the theory to explain all these facts all by itself. We recognize that many facts about X will also involve some other phenomena besides X. So, when we say the theory explains the facts about X, we mean that the theory, in conjunction with relevant background theories pertaining to the other phenomena involved, explains the facts about X. Horwich puts it like this:

In so far as we want to understand truth and the other phenomena, then our task is to explain the relationship between them...We must discover the simplest principles from which they can all be deduced: and simplicity is promoted by the existence of separate theories of each phenomenon. Therefore it is quite proper to explain the properties of truth by conjoining the minimal theory with assumptions from elsewhere...The virtue of minimalism, I claim, is that it provides a theory of truth that is a theory of nothing else, but which is sufficient, in combination with theories of other phenomena, to explain all the facts about truth.8

Are there any constraints on the background theories minimalists are allowed to invoke when attempting to explain the facts about truth on the basis of MT? Although Horwich does not mention any, there must be such constraints. Sure enough, most facts about truth will also involve other phenomena. But many facts about other phenomena will also involve truth. If all facts from theories about other phenomena, including the ones that also involve truth, can be invoked to explain the facts about truth, then the Adequacy Thesis is empty and cannot serve to support the

---

8 Horwich, op. cit., 24-25; cf. also 6-7, 11-12, 20-24.
minimalist view that MT is the right theory of truth — any “theory” of truth can “explain” the facts about truth, when combined with the facts about truth. So the admissible background facts must be restricted to truth-free facts. But there is a complication. Horwich rightly regards logical principles/facts as available for explanatory purposes by default. Now, the principle that every true proposition is true is a logical principle, for it is expressed by an instance of the valid formula ‘(\forall x)((Fx \& Gx) \rightarrow Gx)’. But it is also a principle about truth as well as a principle about propositions. The above restriction would make such principles/facts unavailable when explaining the facts about truth. Let us simply agree, then, that the restriction to truth-free facts from other theories does not extend to principles/facts from logic. (Logical principles involving truth are special in any case: if logic is available for explanatory purposes by default, then logical principles involving truth can be explained without making any use of MT.) Are there further constraints on the admissible background theories? It is difficult to think of any that can be made reasonably precise. Unfortunately, this makes the Adequacy Thesis rather slippery. It is tempting to save the thesis — come what may — by drawing on whatever background “theories” are needed to save the thesis, regardless of their antecedent plausibility.

The passage from Horwich makes clear that, as far as the Adequacy Thesis is concerned, he identifies explainability on the basis of X with deducibility from X. Although such an identification might be considered problematic when applied to causal/scientific explanation, it is appropriate for the special case at hand. According to Minimalism, truth is an insubstantial, quasi-logical property. Part of what this is supposed to mean is that truth is not a genuinely explanatory property. This deflationary character of truth is supposedly exhibited by MT. Hence, it is appropriate, even mandatory, for a minimalist to maintain that whatever explanatory content a fact involving truth plus other phenomena may have, the contribution made
by truth to this content has to be completely reducible to the deflationary theory MT. Requiring that such facts must be deducible from MT (in combination with truth-free facts about the other phenomena involved) should insure that there is nothing more to the truth property involved in such facts than what is covered by deflationary MT. This, then, is another reason why the Adequacy Thesis is crucial to Minimalism. It is needed to support the minimalist claim that the truth property involved in facts about truth is not a substantive property.

Unfortunately, Horwich tends to talk of fact-deducibility rather than sentence-deducibility. Unless tightly constrained by the latter, the former can easily become a loose and very elastic notion. Consider, e.g., the following question: Is the fact (a) that bachelors are unmarried deducible from the fact (b) that unmarried men are unmarried men? Loose answer: Yes, because the fact (a) is just the same fact as (c) that unmarried men are unmarried, and this fact is deducible from the fact (b). Deducibility in this loose sense raises murky issues concerning the identity criteria for facts/propositions and allows one to stretch the notion of deducibility quite considerably (lots of facts are “deducible” modulo more or less defensible assumptions about fact identity). On a strict answer to the above question, (b) is not deducible from (a), because the sentence expressing (a) is not a formal logical consequence of the sentence expressing (b). The derivation requires, as an additional premise, a sentence expressing the identity of (a) with (c), but such a premise is not an instance of a logically valid formula: if (a) is deducible from (b) only modulo a non-logical premise, then (a) is not deducible from (b). Although the strict answer allows for fact-deducibility too, it requires that any such deducibility claim can be cashed-in in terms of sentence-deducibility: a fact named by \([\text{that } S]\) counts as deducible, only if the sentence \(S\) is deducible. Note in particular that necessity claims do not warrant deducibility claims. The intuition that \([S_1 \rightarrow S_2]\) expresses a necessary fact does not underwrite the claim that the fact expressed by \(S_2\) is
deducible, in the strict sense, from the fact expressed by $S_1$. The deducibility claim has to be based on there being a valid logical formula/rule showing that sentences of the form of $S_2$ are formal logical consequences of sentences of the form of $S_1$. To proceed differently means to base loose deducibility claims on potentially murky intuitions of necessity (deducibility is evidence of necessity, not the other way round). Deducibility claims should first and foremost be claims about strict deducibility. Taken in any looser sense, such claims will involve hidden premises or modal intuitions that may well be up for grabs. Construing explainability in terms of some loose notion of deducibility will make the Adequacy Thesis rather slippery. It will make it tempting to save the thesis in times of need by stretching deducibility however far is required for the problem case at hand.

To summarize. The Adequacy Thesis says that MT, when combined with truth-free facts from background theories about other phenomena, provides adequate explanations of all the facts involving truth (keeping in mind that logical principles/facts involving truth are exempted from the restriction to truth-free facts). The notion of explainability is to be understood in such a way that a fact counts as adequately explainable on the basis of MT & X just in case it is deducible from MT & X.

I think that the Adequacy Thesis fails. But, as I have already “foreshadowed,” the thesis is difficult to pin down because it is slippery along two dimensions: admissibility of background theories; deducibility of facts. Let us now take a look at a number of specific cases and see how these issues come up—in varying degrees of severity—when one actually tries to explain/deduce facts involving truth on the basis

---

9 Horwich tends to talk about deducibility of facts as opposed to propositions. As far as I can see, he does so merely because ‘explains’ goes more comfortably with ‘fact’ than with ‘proposition’. At bottom the Adequacy Thesis must be committed to the deducibility of propositions about truth in general. If facts/truths involving truth are explainable/deducible from MT in combination with facts/truths from true theories about other phenomena, then falsehoods involving truth must be
of MT. From now on, I will adopt a convenient notation from Horwich. I will use ‘⟨...⟩’ to abbreviate the noun phrase ‘the proposition that...’, where ‘...’ can be filled with a sentence schema or with an actual sentence.

Consider the fact expressed by:

(1) ⟨Snow is white⟩ is false → ⟨snow is white⟩ is not true.

This fact involves truth as well as falsehood. Evidently, it is not strictly deducible from MT without a background theory of falsehood that supplies an additional premise. One such theory proposed by Horwich is the following: (∀x)(x is false → x is a proposition & x is not true). Of course, MT is now unnecessary for deducing (1). This creates no problem: since (1) is deducible from this theory of falsehood alone, it is trivially deducible from MT plus this theory. But the theory violates the restriction on admissible background theories, for the theory itself involves truth. One could respond that the theory is an explicit definition of falsehood in terms of truth and that explicit definitions involving truth should be exempted from the restriction to truth-free facts—the idea being that explicit definitions are just logical facts in another dress. The claim now would be that (1) is trivially deducible from MT, because it is deducible from a logical fact which is available by default. The logical fact in question is: ⟨snow is white⟩ is not true → ⟨snow is white⟩ is not true, which is claimed to be the same fact as (1) by definition. But the premise that fact (1) is the same as this logical fact, though underwritten by a proposed definition, is not a logical premise. (1) is deducible from the logical fact only modulo the definition; hence, (1) is not deducible from the logical fact. On the present proposal for handling falsehood, the Adequacy Thesis must be interpreted as talking about a notion of deducibility*, where a conclusion is deducible* from X, if it is deducible from X in deducible from MT in combination with falsehoods from false theories about other phenomena. Otherwise, MT would be able to discriminate between truths and falsehoods.
combination with definitions (this is the Frege-Carnap notion of analytic consequence as opposed to logical consequence). To put this the other way round, if we are to deduce facts like (1), as opposed to merely deducing* them, MT must be replaced by a theory MT\(^+\), consisting of MT plus the above theory of falsehood.

Horwich has an alternative proposal for handling falsehood, namely by way of a theory that consists of the collection of axioms of the form:

\[(2) \quad \langle p \rangle \text{ is false } \leftrightarrow \text{not } p,\]

where ‘not’ indicates the logicians external negation operator ‘it is not the case that’.\(^{11}\) Let us grant, if only for the sake of argument, that (2) is truth-free and does not give rise to the difficulties generated by the first theory. However, unlike the first theory, this one is an infinite theory of falsehood. On this proposal, then, the Adequacy Thesis can be sustained only if, in addition to infinite MT, we also accept an infinite theory of falsehood as an admissible background theory. We will encounter this pattern again. Saving the Adequacy Thesis forces the minimalist to make one of two moves (or a combination of them): stretching the notion of deducibility to some notion of deducibility*; or stretching our idea of what counts as an available background theory so that antecedently suspect theories (e.g., infinite theories) count as available.

Consider the question how the minimalist will explain/deduce the following modal fact about truth on the basis of MT:

---

\(^{10}\) See Horwich, op. cit., p. 71.

\(^{11}\) See Horwich, op. cit., pp. 71-72. Horwich presents this as a variant of the proposal discussed above. For the reasons given in the text, I think the two proposals are quite different. Horwich rightly worries that ‘it is not the case that’ amounts to ‘it is not true that’, so that (2) would not constitute any progress over the previous proposal. To allay this worry, he tries to define ‘not’ implicitly without using ‘true’ or close relatives. Let us accept that this is indeed feasible. But there is another problem with (2): if \(\langle p \rangle\) does not exist, then (2)’s right-hand side holds while its left-hand side fails. To repair this, one could replace (2) with: \(\langle p \rangle \text{ is false } \leftrightarrow \langle p \rangle \text{ exists and not } p\).
(3) Necessarily (∃dogs fly) is true → dogs fly).

Invoking a finite background theory, (∃x)(x is a necessary proposition → x is true), violates the restriction to truth-free facts. Unlike the first of the two theories of falsehood, this theory cannot be exempted on the grounds that it is an explicit definition of necessity. Since this theory of necessity is not truth-free, it is tempting to invoke instead the theory expressed by the substitution instances of

(4) (Necessarily p) → p.

which asks us to embrace yet another infinite background theory. But this time the move to the infinite theory does not even work: (3) is not deducible from MT in combination with (4). The desired deduction requires the necessitation of MT, i.e., it requires (3)—among other things—but (3) was what we were trying to deduce. One way out would be to replace MT with its necessitation (or better with: Necessarily (∃p) is true ↔ (∃p) exists and p). Of course, that would mean discarding MT; and we would still need the infinite theory (4) to deduce the propositions hitherto known as the axioms of MT. Horwich makes a different proposal. He suggests, albeit tentatively and in a footnote, that our theory of necessity should contain the following principle: Propositions that are explanatorily fundamental are necessary truths; given that the MT-axioms are explanatorily fundamental, we could derive (3). But the proposed principle is false: natural science contains explanatorily fundamental laws that are not necessary truths. Moreover, the principle is not truth-free. The attempt to make it truth-free by deleting the last word leaves us with an ambiguous principle—

---

12 Could it be exempted on the grounds that it is part of an implicit definition of necessity? Only by trivializing the Adequacy Thesis: there is no limit to principles involving truth that can be proclaimed to function as part of an implicit definition of some notion X.

13 The axioms of MT may be necessary truths, but this does not make (3) deducible from (4) plus MT. Note also that (3) is not deducible from (4) plus MT plus the premise that all axioms of MT are necessary truths; moreover, this premise isn’t truth-free anyway.

14 Cf. Horwich, op. cit., p. 21, n. 5.
‘x is necessary’ could mean ‘x exists necessarily’ or it could mean ‘x is necessarily true’. This problem could be avoided by switching to the schema: If \langle p \rangle is explanatorily fundamental, then it is necessary that p. But the theory expressed by the instances of this schema is not finite. The minimalist is now saddled with the consequence that our theory of explanation has to be infinite too.

John knows that dogs bark, only if \langle dogs bark \rangle is true. To deduce this fact without violating the restriction on admissible background theories, the minimalist has to use the premise: John knows that dogs bark, only if dogs bark. And to deduce the fact that John knows that dogs fly, only if \langle dogs fly \rangle is true, the minimalist has to use the premise: John knows that dogs fly, only if dogs fly; and so on. What theory will supply all these premises? The theory of knowledge—but apparently not by providing a general principle connecting knowledge with truth, for such a principle would not be truth-free. So, because of the Adequacy Thesis, the minimalist is committed to hold that the theory of knowledge supplies each of these premises individually. That is, it has to be construed as yet another infinite theory, one containing a collection of particular axioms expressed by the instances of the schema: S knows that p, only if p. The same will hold for a considerable number of other theories, namely for all theories that are ordinarily regarded as offering general principles connecting their subject matter with truth. All such theories have to be (re)interpreted as consisting, to a large part, of infinite collections of particular axioms rather than finite general principles.\(^{15}\)

Let us consider a somewhat different case. Horwich acknowledges that there is something to be said for the “correspondence intuition” that truths are made true

\(^{15}\) The definition of knowledge contains some clauses in addition to the truth-clause. Once the truth-clause is construed as an infinite collection of axioms, the whole definition has to be construed as an infinite collection of axioms. Otherwise, the connection between the different clauses of the definition would be lost.
by reality. He proposes to account for this intuition by explaining (7) from (5) and (6):

(5) Snow is white;
(6) \langle \text{Snow is white} \rangle \text{ is true;}
(7) \langle \text{Snow is white} \rangle \text{ is true because snow is white.}

Science gives us (5). Given MT, we can then explain (6) from (5). Having done that, we have an explanation of (7), because we have explained (6) on the basis of (5). But (7) is clearly not deducible from these premises. It seems Horwich is trying to use the meta-level claim “If (6) is deducible from (5) given MT, then (7) is deducible” as a premise within the deduction of (7). This cannot work since the antecedent of this meta-level premise is not forthcoming anywhere within the deduction: the fact that X is deducible from Y given Z does not make it deducible that X is deducible from Y given Z. If explainability is here still supposed to be cashed out in terms of deducibility, then Horwich has moved on to some further stretched-out notion of deducibility* that leaves ordinary deducibility far behind.

There is a whole class of facts, namely universal generalizations involving truth, that pose a special challenge for the Adequacy Thesis. After all, MT is merely a collection of particular truths about truth: How will the minimalist account for any universal generalizations involving truth? Objections to the effect that MT-type theories will prove all instances of a given universal generalization but must be too weak to prove the generalization itself have been raised by Tarski (in advance, as it

---

17 Horwich does not say explicitly that the notion of explainability in play here is still connected with deducibility. But if not, what happens to the Adequacy Thesis? Moreover, the same problem arises for explainability itself: that X explains Y does not explain why X explains Y. Incidentally, (7) does not capture the correspondence intuition at all. The intuition says that, if a proposition is true, then it is made true by reality. Unlike (7), this covers falsehoods too. That is, Horwich needs to explain, e.g.: (\langle \text{snow is green} \rangle \text{ is true} \rightarrow (\langle \text{snow is green} \rangle \text{ is true because snow is green}); it is not easy to see how such an explanation would go.
were), and again by Anil Gupta, Scott Soames, and others. Let us see how Horwich addresses this issue. He gives only one example of a minimalist account of a universal generalization involving truth. Presumably, the example is to serve as a model for minimalist explanations of other universal facts about truth. I reproduce Horwich’s account in full:

If one proposition implies another, and the first one is true, then so is the second. Here is a minimalist explanation:

1. Logic provides us with facts like

   \[ \text{dogs bark} \land (\text{dogs bark} \rightarrow \text{pigs fly}) \rightarrow \text{pigs fly}, \]

   that is, with every fact of the form

   \[ [p \land (p \rightarrow q)] \rightarrow q. \]

2. Therefore, given MT, we can go on to explain every fact of the form

   \[ [(p \text{ true}) \land (p \rightarrow q)] \rightarrow (q \text{ is true}). \]

3. But from the nature of implication, we have all instances of

   \[ ((p \text{ implies } q)) \leftrightarrow (p \rightarrow q) \]

4. Therefore we can explain each fact of the form

---

18 It seems Tarski was the first to advocate a minimal theory—but with respect to truth for sentences. Believing that his method of defining truth would not work for “languages of infinite order,” he tentatively advocated adopting as axioms all instances of his schema \( \chi \text{ is Tr } \leftrightarrow p \). Tarski was less than enthusiastic about his own proposal. He observed that the resulting theory “would be a highly incomplete system, which would lack the most important and fruitful general theorems”; see Alfred Tarski, “The Concept of Truth in Formalized Languages,” in Logic, Semantics, Metamathematics, 2nd ed., trans. by J. H. Woodger, ed. by J. Corcoran (Indianapolis: Hackett 1983), p. 257; but see also his subsequent remarks on pp. 258-262. Gupta and Soames both address their objections directly at the first edition of Horwich’s book; see Gupta, “Minimalism,” op. cit., pp. 363-365; Scott Soames, Understanding Truth (New York: Oxford University Press 1999), p. 247. Since Horwich now holds that the complete theory of truth consist of MT plus one general axiom (see note 7), there are some universal generalizations that can be easily handled by his theory. Moreover, as we have seen earlier, generalizations that are instances of logical principles can also be handled. Still, this leaves quite a few generalizations to be accounted for.
\[ [(\langle p \rangle \text{ is true } \& \langle p \rangle \text{ implies } \langle q \rangle) \rightarrow \langle q \rangle \text{ is true}. \]

5. And therefore, given MT, we get each fact of the form
\[ [(\langle p \rangle \text{ is true } \& \langle p \rangle \text{ implies } \langle q \rangle) \rightarrow \langle q \rangle \text{ is true}] \text{ is true.} \]

6. But it is a peculiar property of propositions that any general claim about them—any characterization of all propositions—is made true by the infinite set of particular facts associating that characteristic with each individual proposition.

7. Therefore, in light of 5 and 6, we can explain the general fact:

Every proposition of the form, \[ [(\langle p \rangle \text{ is true } \& \langle p \rangle \text{ implies } \langle q \rangle) \rightarrow \langle q \rangle \text{ is true}], \text{ is true.}^{19} \]

Step 1 makes use of the idea that logical principles/facts are available by default. Step 3 refers to a background theory of the nature of implication. Note in passing that this theory takes the form of yet another infinite collection of propositions and that accounts of generalizations about truth will require recourse to additional infinite background theories. Step 6 seems the most crucial one, the one where we move from particular facts to the universal generalization cited at the very end of the passage. This generalization can be abbreviated as

(7.1) \((\forall x)(\text{IMP}_x \rightarrow T_x)\),

where ‘IMP’ abbreviates the predicate ascribing the form mentioned in the generalization, and ‘T’ abbreviates the truth predicate. Surprisingly, (7.1) is not the generalization whose explanation Horwich announces at the very beginning of the passage. But let us suppress this for now and let us see whether the account succeeds with respect to (7.1).

\(^{19}\)See Horwich, op. cit., 21-22. I have made a change in step 3 where Horwich has only ‘\((\langle p \rangle \text{ implies } \langle q \rangle) \rightarrow (p \rightarrow q)\)’ which must be a typographical error.
Horwich’s account has the following curious feature: it operates on two levels at once. On one level, the meta-level, we are presented with an argument consisting of steps 1-7 and culminating in a conclusion about (7.1), namely the conclusion that (7.1) is explainable. On the face of it, this meta-level argument cannot actually be the explanation of (7.1). The actual explanation of (7.1), its object-level deduction, must be the one indicated by the indented lines. Of course, if the meta-argument 1-7 is sound, then (7.1) is explainable/deducible. But it seems that to check whether the argument is indeed sound, we have to see whether (7.1) is deducible at the object-level, i.e., from premises like the ones indicated by the indented lines (I say “indicated” because, unlike (7.1), all but one of the lines indented in 1-5 are mere schemata, truth-valueless proxies for actual premises). Right away, we encounter an obstacle: What about the crucial step 6—it has no indented line? Which level does it belong to?

Let us first try to interpret 6 as supplying a premise to be inserted into the object-level (indented) deduction. Let us use ‘a₁’, ‘a₂’,... as abbreviations of propositional names of the form ‘⟨⟨p⟩ is true & ⟨p⟩ implies ⟨q⟩⟩ → ⟨⟨q⟩ is true⟩’, which is just the form IMP. So ‘Ta₁’, ‘Ta₂’,... attribute truth to each proposition of form IMP, but without saying that it is of form IMP. According to step 5, each of these truth attributions is deducible at the indented level. One might think we could deduce (7.1), if 6 were to supply us with the premise:

\[(6.1) \quad ([Ta_1 & Ta_2 & \ldots) & (\forall x)(IMP \rightarrow x = a_1 \lor x = a_2 \lor \ldots)] \rightarrow (\forall x)(IMP \rightarrow T_x).\]

The second conjunct of the antecedent says that a₁, a₂,... are all the propositions of form IMP. The consequent is the desired (7.1). To detach it, we need the antecedent. But the second conjunct of the antecedent is entirely unsupported by the previous

---

20 This feature, as well as step 6, are new to the second edition.
indented premises. Moreover, to get at the first conjunct, we would need to complete infinitely many deductions of the style indicated by the indented lines in 1-5; and we would need to assemble the results into an infinite conjunction. There is no rule of logic that allows us to do that. In addition, (6.1) cannot be used as a premise anyway. It is not an instance of a well-formed formula because it is infinite. The last problem could be circumvented, if 6 were taken to insert either one of the following finite premises into the deduction of (7.1) as an indented line:

(6.2) If all the propositions that are the instantiations of a universal generalization over propositions are true, then the universal generalization is true;

(6.3) The set containing all the instantiations of a universal generalization over propositions implies the universal generalization.

But neither of these allows us to deduce (7.1) from the available premises. Their application would require an infinite premise, saying of each one of ‘Ta₁’, ‘Ta₂’,... that it is an instantiation of (7.1), and saying that these are all the instantiations of (7.1). Again, no such claim is deducible from the prior indented premises. Moreover, version (6.2) is itself a universal generalization involving truth. If it is required for deducing universal generalizations involving truth, all such deductions will be hopelessly circular. Version (6.3) is not innocuous either. A set of premises implies a conclusion only if the corresponding conditional is true—just ‘true’, if ‘imply’ refers to material implication, ‘necessarily true’, if it refers to necessary implication. The conditional corresponding to (6.3) would have to have an infinite formula as its antecedent. Since such conditionals are not well-formed, what (6.3) says is not true.

In a later passage of his book Horwich tries to clarify the status of step 6. He points to two prima facie problems for the minimalist attempt to explain universal generalizations involving truth.²¹ First, there is no logically valid rule enabling us to

²¹ Cf. Horwich, op. cit., p. 137.
assemble all the required premises. I think Horwich has in mind that there is no logical rule for deducing infinite formulas. This is so not only because infinite formulas are not well-formed, but also because no rule enables us to derive any conclusion (not even a finite one) from all the premises of an infinite premise set: if X is deducible from a set \( \exists \), then X is deducible from a finite subset of \( \exists \). Second, the collection of instantiations of a universal generalization does not entail the generalization itself. Although Horwich seems to regard these as two aspects of the same problem, the second is quite different, for it has nothing in particular to do with infinity. Consider the set of premises saying of each of the nine solar planets that it has property F. There could have been an additional planet that failed to be F; hence, there is a possible world, different from ours, in which all the premises are true but the conclusion that all solar planets are F is false. Horwich thinks the second problem does not really arise when the objects in question are propositions:

It seems to me that in the present case, where the topic is propositions, we can find a solution to this problem. For it is plausible to suppose that there is a truth-preserving rule of inference that will take us from a set of premises attributing to each proposition some property, F, to the conclusion that all propositions have F. No doubt this rule is not logically valid, for its reliability hinges not merely on the meanings of the logical constants, but also on the nature of propositions. But it is a principle we do find plausible.22

If we are to use Horwich’s rule to infer (7.1), we must identify the property F that (7.1) attributes to all propositions. It is the conditional property expressed by \( \text{IMP} \rightarrow T \). According to step 5, we can get \( \text{T}_a, \text{T}_b, \ldots \), where \( a, b, \ldots \) are the propositions of form IMP. From these we can deduce \( \text{IMP}_a \rightarrow \text{T}_a \), \( \text{IMP}_b \rightarrow \text{T}_b \), \ldots, where \( a, b, \ldots \) are again propositions of form IMP. Each of these premises
attributes the conditional property to a proposition of form IMP. But this is of no use. In order for Horwich’s rule to apply, each of its input premises has to attribute the conditional property to a proposition, period. No such premises are made available.\textsuperscript{23}

It seems, then, that the inference rule should be described as a rule that will take us from a set of premises attributing to each proposition of form \( f \) some property, \( F \), to the conclusion that all propositions of form \( f \) have \( F \). Let us take this as a description of Horwich’s rule.

Why does Horwich think the second problem is held at bay due to the special nature of propositions? It must be because he thinks that propositions—unlike, say, planets—obey the following principle: If a proposition exists in any possible world, then it exists in every possible world. But this principle holds only on some (e.g., Fregean) conceptions of the nature of propositions. According to some popular alternatives, there are so-called singular propositions. They are said to have individual objects as their constituents; consequently, their existence depends on the existence of these objects. So, any world in which there is an object, say, a planet, that does not exist in our world is a world in which there is also a proposition that does not exist in our world; and such a world may falsify Horwich’s rule. The upshot is a significant weakening of the Adequacy Thesis. It turns out that the thesis holds only \textit{modulo} a specific, and contentious, theory of the nature of propositions, according to which there are no singular propositions. Moreover, according to Aristotelians, a property (universal, concept, etc.) exists only if there is at least one object that exemplifies it. On this view, the existence of a proposition containing a property will depend on there being an object exemplifying this property. So the Adequacy Thesis holds only \textit{modulo} a thoroughly Platonistic conception of

\textsuperscript{22} Horwich, op. cit., 137, some of the italics are mine.

\textsuperscript{23} This difficulty would remain, even if there were only finitely many propositions of form IMP. What we need, according to the rule, are premises attributing the conditional property to each proposition, including those that are not of form IMP: the problem is that ‘\( Ta_1 \)’, ‘\( Ta_2 \)',... are not instantiations of (7.1).
propositional constituents.  

In the passage quoted above, Horwich seems to be telling us that step 6 is not intended as a premise to be inserted into the object-level (indented) deduction. Rather, 6 is a rule of inference—or better, 6 expresses a principle, namely (6.2), which underwrites a rule of inference, namely the one hinted at in the quoted passage. The rule is supposed to license the transition from the premises available under step 5 to the conclusion (7.1). But the rule is not spelled out. It is only circumscribed in terms of what it ought to do. It is supposed to operate on all premises attributing a property $F$ to propositions of form $f$, even in the absence of any premise saying that $a_1, a_2, \ldots$ are all the propositions of form $f$. The characterization of the rule seems designed to avoid the need for assembling infinitely many premises into an infinite conjunction, as well as the need for assembling an infinite premise to the effect that the conjunction contains all the relevant premises, but Horwich does not explain how the envisioned inference rule might work in the absence of such “assemblings.” Consequently, it is hard to see how the alleged “rule” can be coherently conceived of as a rule at all. It seems Horwich’s claim that there is such a rule rests on the intuition that, given the appropriate conception of propositions, principle (6.2) is a necessary truth. But the mere presence of an (alleged) necessary truth is not sufficient for underwriting the claim that there is a corresponding rule of inference. 

---

24 I should point out that (7.1) has a special and potentially misleading feature. Since IMP is a form of necessary propositions, every proposition of form IMP is true in every world in which it exists. So no world can serve up a counterexample to (7.1). But this is entirely due to the special generalization originally chosen by Horwich. Generalizations about propositions of other forms will give rise to the problem mentioned in the text. Note that something similar holds for planets. If $E$ is a necessary property of planets, then no world can offer a counterexample to the claim that all planets are $E$; this works only for necessary properties.

25 Even if one believes that “Snow is white $\rightarrow$ numbers (God, possible worlds) exist” expresses a necessary truth, given the appropriate conception of numbers (God, possible worlds), one should admit that the corresponding inference is not a good one. Horwich objects to theories other than MT on the grounds that they fail to explain the axioms of MT (see, e.g., op. cit., pp. 11-12). But if an
Horwich’s rule flouts the principle that X can be inferred from an infinite premise set if only if X can be inferred from a finite subset of : there is no finite conditional corresponding to the rule. Moreover, the rule is “object specific”; that is, it works only for certain objects, propositions—worse, it works only modulo a specific theory of propositions. For all these reasons, Horwich’s rule cannot be a valid rule of deduction. Horwich himself admits this, of course, when he points out that the rule is not logically valid. But does this not mean that the Adequacy Thesis fails? By my lights, it does. However, it is always possible to claim that conclusions not deducible from by valid rules of deduction come out as deducible* from by valid* rules of deduction*.

One could try to elevate such a claim above the level of bare assertion by way of a very non-standard notion of deducibility—call it meta-deducibility—characterized along the following lines: If there is a meta-level argument showing that every fact attributing a property F to a proposition of form f is deducible from , then the proposition that all propositions of form f have F is to count as meta-deducible from . A fact involving truth would then count as explainable on the basis of MT and background theory X, if it is deducible or meta-deducible from MT & X. This notion of meta-deducibility does not fit well with the rule hinted at by Horwich. However, since the notion has the very unusual feature that it would allow one to “establish” an object-level conclusion via a meta-level argument, it does fit with Horwich’s meta-level argument, 1-7, for the explainability of (7.1)—Maybe this argument should after all be regarded as itself constituting the explanation of (7.1)? This would mean that universal facts about truth are not explainable, in the old sense, on the basis of MT & X. Instead, universal facts attributing truth to all propositions

intuition of the necessity of X→Y were sufficient to warrant the claim that X explains Y, then every theory of truth would explain the axioms of MT since they are necessary truths.
of form $f$ are explainable, in a new sense, by the fact that each particular fact attributing truth to a proposition of form $f$ is explainable/deducible, in the old sense, on the basis of MT & X. It is hard to see how this goes beyond the bare assertion that the general fact is explainable because the particular facts are explainable—the very claim that was at issue to begin with. Moreover, this new notion of explanation is rather puzzling. Ordinarily, one thinks that facts about objects of kind K are explained by other facts about objects of kind K, L, and M. The claim that certain facts about Ks are explained by the fact that we can explain other facts about Ks suggests that the facts in question are not regarded as real facts.

A minimalist may want to respond to all this: “Why be so conservative? Why not accept a non-standard notion of deducibility*?” Let us grant, for the sake of argument, that some appropriate non-standard notion(s) can be made sense of. The result must be that the Adequacy Thesis is far weaker than originally advertised. I remarked earlier that the thesis plays a crucial role in the minimalist best-explanation argument for the exclusivity of MT, where it is argued that other theories of truth ought to be rejected because they fail to provide adequate explanations of the facts about truth. For such an argument to work, there has to be a shared standard of adequate explainability. Deducibility from the proposed truth theory (plus background theories) could serve as such a standard. Given this standard, one half of the minimalist best-explanation argument seems indeed correct: other truth theories do not enable us to deduce all the facts about truth. However, neither does MT. To save the Adequacy Thesis, the minimalist trades deducibility for some extended notion of deducibility*. But now the common standard of adequate explanation is discarded and the best-explanation argument loses all force. Other truth theories will be able to provide adequate explanations of the facts about truth, provided they can

---

26 Of course, meta-deducibility still depends on the specific conception of the nature of propositions remarked on above. Tarski considers a similar notion—applied to numbers rather than propositions—in his “The Concept of Logical Consequence”, in Tarski, op. cit., p. 411.
select some notions of “deducibility” appropriately extended to serve their needs. I also remarked that the Adequacy Thesis, construed in terms of deducibility, would help ensure that the facts about truth are completely reducible to MT (and truth-free background theories). This would lend some support to the minimalist claim that the facts about truth do not involve a truth property more substantial than the one covered by deflationary MT. But once deducibility is traded in for some extended notion of deducibility*, the Adequacy Thesis lends no support to this claim. Why not claim, on the contrary, that truth must be a substantial property because universal generalizations about truth are not deducible from MT?

Similar considerations apply with respect to the issue of admissible background theories. The Adequacy Thesis can be defended only relative to a very specific background theory about the nature of propositions; moreover, background theories that are somehow enmeshed with truth have to be reconstrued as infinite theories. Such specific and contentious commitments help save the Adequacy Thesis only by weakening it.

I want to return briefly to a curious feature of Horwich’s account of universal facts about truth that I have set aside earlier. At the outset, Horwich promises an explanation of (Imp), but his account ends with (7.1):

(Imp) For all propositions \( x, y \): if \( x \) is true, and \( x \) implies \( y \), then \( y \) is true;

(7.1) For all proposition \( x \): if \( x \) is of the form, \( \langle \langle p \rangle \text{ is true } \& \langle p \rangle \text{ implies } \langle q \rangle \rangle \rightarrow \langle \langle q \rangle \text{ is true} \rangle \), then \( x \) is true.\(^{27}\)

\(^{27}\) It seems that (7.1) is misstated. Since ‘(...)’ abbreviates ‘the proposition that...’, (7.1) comes out as “Every proposition of the form that the proposition that \([\langle p \rangle \text{ is true } \& \langle p \rangle \text{ implies } \langle q \rangle \] \rightarrow \langle \langle q \rangle \text{ is true} \] is true,” which does not make any sense. The outermost ‘(...)’ in (7.1) should be replaced by quotes. This fits in with what Horwich says about form-talk elsewhere. Note that \( \langle \text{dogs fly } \rightarrow \text{ dogs fly} \rangle \) is not of the form ‘\( \langle p \rightarrow p \rangle \)’; rather, it has the form ‘\( p \rightarrow p \)’. The former is not a form of a proposition at all because it is not the form of a sentence expressing a proposition; it is the form of a noun-phrase referring to a proposition; cf. Horwich, op. cit., p. 123.
On the face of it, these generalizations differ in content as well as in form. This shift from one generalization to another will be a general trait of all explanations modeled on Horwich’s account. When we ask the minimalist to explain, say, the generalization that a proposition is known only if it is true, the last line of his argument will offer us instead the generalization that every proposition of the form ‘If \( \langle p \rangle \) is known then \( \langle p \rangle \) is true’ is true. This raises the objection that Minimalism is unable to explain ordinary general facts about truth; instead, it offers us general form-facts about truth. It seems the minimalist will have to respond that (Imp) and (7.1) express the same fact. To put it more generally and in the material mode, he will have to maintain that ordinary general facts about truth are really form-facts about truth. This response further weakens the Adequacy Thesis. For it turns out that the thesis holds only modulo a contentious claim about fact identity. The claim would have to be backed-up with some theory of forms that applies to facts and propositions—such a theory remains to be spelled out (and it is sure to be contentious too).

---

28 At times Horwich construes form-talk linguistically, so that “Every proposition of the form ‘...p...’ is true” comes out as “Every proposition expressed by a sentence of the form ‘...p...’ is true”; cf. Horwich, op. cit., p. 123. On this interpretation, form-facts like (7.1) would come out as disguised linguistic facts and the thesis that all ordinary general facts are form-facts would be quite untenable. Horwich may ultimately prefer a non-linguistic interpretation of form-talk, on which the thesis might be somewhat less implausible. But there is little by way of a non-linguistic theory of forms. On pp. 17-20 he tries to construe propositional forms/structures as functions—unsuccessfully by my lights. He holds that the propositional form/structure

\[
(\text{E*}) \quad \langle \langle p \rangle \rangle \text{ is true iff } p
\]

is a function from propositions to propositions. If so, it would have to make sense to say that, given a proposition as argument, there is a proposition which is the value of the function (E*). That is, it would have to make sense to say that, for every proposition \( \chi \), there is a proposition \( \xi \), such that \( \xi = \langle \langle \chi \rangle \rangle \text{ is true iff } \chi \). But this does not make sense. Take \( \chi = \text{the proposition that snow is white} \); we then have \( \xi = \text{the proposition that the proposition that the proposition that snow is white is true is true iff the proposition that snow is white} \).