

Minimalism, Deflationism, Pragmatism, Pluralism

Crispin Wright

1 Minimalism and Deflationism: Overview

Deflationists have offered views about truth differing significantly in detail. But they characteristically maintain that as far as philosophy is concerned, there is nothing to say about truth that is not captured by a suitably generalized form of one (or both) of the following two schemata:

(ES) It is true that P iff P .

(DS) “ P ” is true iff P .

And they maintain that this point in turn entails deflation—that the traditional metaphysical debates about truth, as well as more recent ones, are about nothing substantial.

It is worth noting that these are separable claims. Someone could allow that the two schemata—the Equivalence Schema, (ES), for propositions and the Disquotational Schema, (DS), for sentences—are each a priori correct¹ and (together) somehow fully encapsulate all proper uses of the truth predicate, without conceding that (it follows therefrom that) truth is somehow not a proper object of further philosophical enquiry, that no further metaphysical or semantic issues arise. Conversely, someone broadly in agreement with the antimetaphysical spirit of deflationism might hold that a correct characterization of the use of the truth predicate demands something more complicated than the two schemata.

The minimalist view about truth that I here defend rejects each of these deflationist claims, contending both that the two schemata are insufficient to capture all that should properly be reckoned as belonging to the concept of truth and that the antimetaphysical message of deflationism,

globally applied, represents a philosophical mistake.² Still, there are points of affinity between minimalism and deflationism. Minimalism agrees that, as far as the *conceptual* analysis of truth is concerned, matters should proceed by reference to set of basic a priori principles in which (ES) and (DS) are preeminent candidates for inclusion, and agrees too that aptitude for truth and falsity goes with surface assertoric content and is not the kind of deep property that, for instance, expressivist views about moral judgement standardly take it to be. However, minimalism rejects the idea that the analysis of the concept of truth exhausts the philosophy of truth: rather, even if the *concept* may be fully characterized by reference to certain basic a priori principles concerning it, the question of which *property* or *properties* of propositions, or sentences, realize the concept can still sensibly be raised for every discourse in which truth has application. Not that an answer to this question has necessarily to provide an identification of truth in the form “ x is true iff x is F .” Minimalism only requires that each discourse that deals in truth-apt claims is associated with such a property whose character need not be fully determinable just from the list of basic principles serving to characterize the concept but which, relative to the discourse in question, serves as truth by dint of satisfying those principles. The fuller characterization of this property will depend on specific features of the particular discourse, and it will ultimately depend on these features whether or not the relevant truth property can be explicitly identified by, for instance, a biconditional of the type above.³

Minimalism thus incorporates a potential *pluralism* about truth, in the specific sense that what property serves as truth may vary from discourse to discourse. And it is this point which allows it to provide hospitality for the discussion of metaphysical—realist or antirealist—ideas that have fuelled those other traditional conceptions of truth that deflationists sought to undermine from the start. This potential pluralism is itself in opposition to the more traditional positions, insofar as they claim to uncover *the* universal nature of truth, something common to all truth-apt discourse. But it can still allow that some regions of discourse may be subject to a truth property congenial to broadly realist thinking about them, while in other regions the character of the truth property may be more congenial to antirealism.

All this may seem to suggest that the key difference between minimalism and deflationism resides in the fact that while the latter concedes the significance of the predicate "true," and hence grants that there is a discussible *concept* of truth, it holds—in contrast to minimalism—that there is no *property* of truth: no property that all truths in a given area have in common.⁴ This view of the matter would be encouraged by some of the literature in the field, but it is not the happiest way of putting the differences. For once the currency of a concept of truth is granted, it ought to be allowed that all truths have at least the following property in common: the property of falling under this concept.⁵ No doubt this move may not illustrate the most natural or fruitful way of conceiving the relationship between concepts and their associated properties in general. But, for all that, it would be misleading to suggest that (most) deflationists would embrace the view that "Coal is black" and "Snow is white" have no more in common than do coal and snow.

The real distinction, then, between minimalism and deflationism in respect of the issue whether truth is a property is not that deflationism cannot consistently allow that it is, but rather that minimalism allows more: precisely, that the character of the property may not be transparent from the analysis of the concept. So in this respect there is a rough analogy with the relationship—to have recourse to a tired but useful example—between the concept of water and the property (that of being composed of H₂O molecules, I suppose) that it denotes. Not that minimalism suggests that it should comparably be an a posteriori matter what property truth (locally) is. It will be a matter for further *conceptual* reflection—of a sort I will try to illustrate in the sequel—what (kind of) property best fulfills (locally) the role circumscribed by the concept. (That is why the water analogy is imperfect.)

This kind of substantial distinction between a concept, *F*, and the property it denotes, being *F*, is called for whenever we stand in need of some sort of general explanation of a characteristic of items that are *F* that cannot be elicited solely from materials directly implicated in those items' falling under the concept in question.⁶ To take a simple instance, suppose, to pursue the tired example, that the concept of water is a natural-kind concept after the fashion of Putnam's well-known paradigm: that it is, e.g., given as the concept of that colorless, odorless,

tasteless liquid that is typically found in lakes and rivers, assuages thirst, and so on. If we allow that it makes good sense to ask *why* water typically presents with the surface features mentioned in its concept, we accept that there is a good explanatory question that cannot, obviously, be answered by appeal to water's falling under its concept, since we are asking for an explanation of the very features involved in its so doing. To allow the legitimacy of the question thus involves conceiving of whatever makes water what it is as distanced from the characteristics presented in its concept—as something that can potentially be invoked in explaining their habit of co-occurrence. But what makes water what it is is just its having the property of being water.

Now, it is plausible enough that there are no such explanations that might be given by appeal to the “thin” truth property that we envisaged the deflationist as admitting—the property of *falling under the concept of truth*—that we could not equally well give by appeal to the concept of truth itself. What the minimalist should claim, accordingly, in contradistinction to the deflationist, is that there *are* certain legitimate explanatory burdens that can be discharged only if we appeal to a property (or properties) of truth conceived in a more substantial sense of “property.”⁷ And note that this claim can be true—in contrast with the situation of the kinds of explanation that might be given by appeal to the property of being water—even if truth, locally or globally, admits of no naturalistic (physicalistic) reduction. (It all depends on whether the things that need explaining are themselves so reducible.) As we shall see in due course, however, the minimalist's argument has no connection with the question of the feasibility of any such reduction.⁸

2 The Inflationary Argument

The inflationary argument is to the effect that the legitimacy of thinking of truth, in any particular discourse, as substantial in a fashion deflationism cannot accept, is already guaranteed by the very principles characterizing the concept of truth to which deflationism gives centre stage—at least when they are taken in conjunction with certain further uncontroversial principles. Thus minimalism does not just go beyond what deflationism allows but contends in addition that deflationism is incoherent: that, in

coupling the thesis that (ES) and/or (DS) yield(s) a complete account of truth with the contention that truth is a property only in the etiolated sense we have just reviewed, its proponents withdraw with one hand what they just tabled with the other.

We begin with the observation that truth-apt contents, or sentences expressing such contents, demand a distinction between circumstances under which asserting them is warranted and those under which it is not. And competent thought and talk requires an ability to tell the difference: I need to be able to tell which assertions I am warranted in making in a given state of information and which I am not. So if I am warranted in asserting *P*, that fact will be recognizable to me, and I will thereby be warranted in claiming that I am so warranted. Conversely, if I am warranted in thinking that the assertion of *P* is warranted, I will be beyond relevant—that is, epistemic—reproach if I go on to assert it. But that is to say that I will be warranted in doing so. We accordingly obtain:

There is warrant for thinking that [it is warrantably assertible that *P*] iff there is warrant for thinking that [*P*].

Given the Equivalence Schema, this will in turn yield:

There is warrant for thinking that [it is warrantably assertible that *P*] iff there is warrant for thinking that [it is true that *P*].

And now, since warranted assertibility is, in a perfectly trivial sense, a normative property—a property possession or lack of which determines which assertions are acceptable and which are not—it follows that truth is too. For by the above equivalence, to be warranted in thinking that *P* is true has exactly the same normative payload as being warranted in thinking that it is warrantably assertible. Moreover, our finding is that truth, as characterized by the schemata, and warranted assertibility *co-incide in positive normative force*.

That is hardly a startling finding. But the relevant point is not the result itself but its provenance: that truth's being normative in the fashion noted is not merely plausible anyway but is a consequence of what ought to be uncontroversial considerations about the concept of assertibility and a central tenet of deflationism: the conceptual necessity of the Equivalence Schema. However, given only the further assumption that any

P apt for truth has a significant negation that is likewise apt for truth, the Equivalence Schema will also entail any instance of the following *Negation Equivalence*:⁹

(NE) It is true that [not- P] iff it is not true that [P].

And this shows that, coincident in positive normative force though they may be, we cannot in general *identify* truth and warrant. For most propositions about most subject matters allow of neutral states of information: states of information in which there are neither warrants for asserting P nor for asserting its negation. In any such case, an invalid schema results if we substitute “is warrantably assertible” for “is true” in (NE). More specifically, if the propositions that make up the substitution class for P allow in principle of neutral states of information, the following conditional is *not* valid:

It is warrantably assertible that [not- P] if it is not warrantably assertible that [P].

Thus, we can already conclude from (NE), and hence from (ES), that truth and warrantable assertibility, even if coinciding in positive normative force, are *potentially divergent in extension*.¹⁰

It is an immediate consequence of this observation that for any assertoric practice that allows the definition, on the contents of the moves it permits, of a truth property satisfying (ES)—that is, for any assertoric practice whatever—there must be a further kind of distinction between circumstances in which making these moves is in good standing and circumstances in which it is not—a distinction that need not coincide with the distinction between circumstances in which such a move can warrantably be made and those in which it cannot. The concept of truth as characterized by (ES) precisely calls for a norm—a way an assertion may be in good standing—which warrant is essentially warrant to suppose satisfied but which, because of the point about potential extensional divergence, may nevertheless not be satisfied when an assertion is warranted (or may be satisfied when it is not). And a fully intelligent participation in such practices will involve grasp that they essentially involve submission to a standard the meeting of which need not just be a matter of possessing warrants for the claim that it is met.

Minimalism now claims that these facts about assertoric practices stand in need of explanation. In particular, it maintains that it needs to be explained what this further norm of correctness amounts to in such a way that it becomes clear how it and warranted assertibility, although potentially divergent in extension, coincide in normative force: how it can be that warrant is essentially warrant to think that this other norm is satisfied when there is no guarantee that they are always *co*-satisfied. And such an explanation, it is contended, while it will have to do much more than this, must at least begin by finding something for the truth of a proposition to consist in, a property that it can intelligibly have although there may currently be no reason to suppose that it has it, or may intelligibly lack even though there is reason to think that it has it. Warrant can then be required to be whatever gives a (defeasible) reason to think that a proposition has that property.

The deflationist account of truth would appear, however, to have no resources to give such an explanation. For all we can elicit from the Equivalence Schema is the *problem*. The point of the inflationary argument is precisely that the basic principles on which deflationism builds its account spawn the concept of a norm—a way a proposition can be in good or bad standing, as I put it a moment ago—that contrasts with its current evidential status. But these principles keep silence when the question is raised, What does the satisfaction or nonsatisfaction of this new norm consist in, and how can it fail to be a substantial property?

So at any rate the inflationary argument contends. But the deflationist is likely to believe that she has a good response. “There is no silence on the point,” she will reply. “On the contrary, my theory is very explicit about what the satisfaction of your ‘norm’ consists in. The proposition that snow is white satisfies it just if snow is white; the proposition that grass is green satisfies it just if grass is green, the proposition that there is no life on Mars satisfies it just if there is no life on Mars. . . .” However, this response is, of course, to no avail unless we *already* understand the difference between the proposition, e.g., that there is no life on Mars and the proposition that that proposition is warranted. And clearly this distinction cannot be recovered from any contrast between the circumstances under which the two propositions are *respectively* warranted, since—as in effect noted right at the start of the argument—there is none.

The difference between them resides, rather, precisely in a difference in correctness conditions of another sort (whisper: *truth* conditions): in order to understand the contrast between the two propositions, I precisely have to understand that the former is in principle hostage to a kind of failure that can occur even when it is warranted, and that will not then affect the latter. So the debate is rapidly brought back to the point before the deflationist made her putative “good response,” with the minimalist charging her to explain (i) how the relevant contrast can so much as exist unless there is something substantial in which such failure—or more happily, success—consists, and (ii) how a *grasp* of the contrast can anywhere be possible unless we are familiar with a (perhaps local) property that behaves as the concept characterized by the basic principles demands.

The kind of move we just envisaged a deflationist making is, of course, pure deflationist stock-in-trade. Supporters of deflationism characteristically view the whole debate as turning on whether it can be shown that all legitimate uses of the word “true” can somehow be explained on the basis of the Equivalence Schema (and/or the Disquotational Schema) together with a repertoire of contexts free of “true” and its cognates, and they put all their effort and (often considerable) ingenuity into the attempt to show that these uses can be so explained.¹¹ But success in this project is entirely beside the point if the contents of the relevant “true”-free contexts, to which deflationists simply help themselves, cannot be explicated by construing them merely as subject to norms of assertibility but demand an additional truthlike constraint. Deducing some aspect of our use of the predicate “true” by appeal just to the Equivalence Schema and certain “true”-free contexts cannot just be assumed to have reductive significance without further ado. The initial position in the debate is one in which nothing yet stands against the opposed thought that, instead of reading the Equivalence Schema from left to right, as if to eliminate the truth property, we should read it from right to left, as highlighting the fact that, implicit in any content in the range of “*P*,” there is already a tacit invocation of the norm of truth. Deflationism needs to get to grips with this reading: to make a case that no implicit prior grasp of the concept of truth, nor implicit reference to a property that the concept denotes, lurks buried in the materials to which its “explanations” appeal. The thrust of

the inflationary argument is that no such convincing case can be made—that whether or not we can somehow eliminate or otherwise “deflate” the *word*, a corresponding property, and its contrast with assertibility, is part and parcel of assertoric content itself.¹²

3 Pragmatism and Pluralism (I): Peirce and Putnam

Let me now be a little more explicit about how minimalism opens up prospects for a pluralistic conception of truth. Above, I spoke approvingly of the idea, of which the deflationist proposals can be seen as one example, that as far as the conceptual analysis of truth is concerned, matters should proceed by reference to a set of basic a priori principles variously configuring or bearing on the concept. Many philosophers, from Frege to Davidson, have, of course, doubted whether truth allows of any illuminating philosophical analysis. But their skepticism has been driven largely by the traditional notion that success in this project would have to consist in the provision of a satisfactory necessary-and-sufficient-conditions analysis of the concept, and there is clearly some scope for relaxation of that model. After all, such a necessary-and-sufficient-conditions analysis, even if it could be provided, would only culminate in one particular a priori—presumably, conceptually necessary—claim. Why should not other such claims—even if not biconditional or identity claims—provide illumination of essentially the same kind? To be sure, if one wants a priori conceptual clarity about what truth—or beauty, or goodness, etc.—is, then the natural target is an identity (or a biconditional). But perhaps the sought-for reflective illumination can be equally well—if less directly—provided by the assembly of a body of conceptual truths that, without providing any reductive account, nevertheless collectively constrain and locate the target concept and sufficiently characterize some of its relations with other concepts and its role and purposes.

What should such principles be for the case of truth? The method here should be initially to compile a list, including anything that chimes with ordinary a priori thinking about truth—what I shall call a *platitude*—and later to scrutinize more rigorously for deductive articulation and for whether candidates do indeed have the right kind of conceptual plausibility. So we might begin by including, for instance,

- the transparency of truth—that to assert is to present as true and, more generally, that any attitude to a proposition is an attitude to its truth—that to believe, doubt, or fear, for example, that *P* is to believe, doubt, or fear that *P* is true. (*Transparency*)
- the epistemic opacity of truth—incorporating a variety of weaker and stronger principles: that a thinker may be so situated that a particular truth is beyond her ken, that some truths may never be known, that some truths may be unknowable in principle, etc.¹³ (*Opacity*)
- the conservation of truth-aptitude under embedding: aptitude for truth is preserved under a variety of operations—in particular, truth-apt propositions have negations, conjunctions, disjunctions, etc., which are likewise truth-apt. (*Embedding*)
- the Correspondence Platitude—for a proposition to be true is for it to correspond to reality, accurately reflect how matters stand, “tell it like it is,” etc. (*Correspondence*)
- the contrast of truth with justification—a proposition may be true without being justified, and vice versa. (*Contrast*)
- the timelessness of truth—if a proposition is ever true, then it always is, so that whatever may, at any particular time, be truly asserted may—perhaps by appropriate transformations of mood, or tense—be truly asserted at any time. (*Stability*)
- that truth is absolute—there is, strictly, no such thing as a proposition’s being more or less true; propositions are completely true if true at all. (*Absoluteness*)

The list might be enlarged, and some of these principles may anyway seem controversial. Moreover, it can be argued that the Equivalence Schema underlies not merely the first of the platitudes listed—Transparency—but the Correspondence Platitude¹⁴ and, as we have seen in discussion of deflationism, the Contrast Platitude as well.

There’s much to be said about this general approach to conceptual analysis, and many hard and interesting questions arise, not least, of course, about the epistemological provenance of the requisite basic platitudes. But such questions arise on *any* conception of philosophical analysis, which must always take for granted our ability to recognize basic truths holding a priori of concepts in which we are interested.

Let us call an account based on the accumulation and theoretical organization of a set of such platitudes concerning a particular concept an *analytical theory* of the concept in question.¹⁵ Then the provision of

an analytical theory of truth in particular opens up possibilities for a principled pluralism in the following specific way: in different regions of thought and discourse, the theory may hold good a priori of—may be satisfied by—different properties. If this is so, then always provided the network of platitudes integrated into the theory is sufficiently comprehensive, we should not scruple to say that truth may consist in different things in different such areas: in the possession of one property in one area, and in that of a different property in another. For there will be nothing in the idea of truth that is not accommodated by the analytical theory, and thus no more to a concept's presenting a truth property than its validating the ingredient platitudes. In brief, the *unity* in the concept of truth will be supplied by the analytical theory, and the *pluralism* will be underwritten by the fact that the principles composing that theory admit of *variable collective realization*.

An illuminating case study for these ideas is provided by pragmatist conceptions of truth. In a very famous passage, C. S. Peirce writes,

Different minds may set out with the most antagonistic views, but the progress of investigations carries them by a force outside themselves to one and the same conclusion. This activity of thought by which we are carried, not where we wish but to a fore-ordained goal, is like the operation of destiny. No modification of the point of view taken, no selection of other facts to study, no natural bent of mind even, can enable a man to escape the predestinate opinion. This great law is embodied in the conception of truth and reality. *The opinion which is fated to be ultimately agreed by all who investigate is what we mean by the truth*, and the object represented in this opinion is the real. [My italics]¹⁶

Here Peirce seemingly believes in a predestined march towards a stable scientific consensus among “all who investigate,” but the received understanding of the “Peircean” view, whether historically faithful or not, has come to be, rather, that the true propositions are those on which investigators *would* agree if—which may well not be so—it were possible to pursue enquiry to some kind of ideal limit; that

P is true if and only if, were epistemically ideal conditions to obtain, *P* would be believed by anyone who investigated it.

An equally famous passage in Hilary Putnam's *Reason, Truth, and History* has regularly been interpreted as advancing the same proposal. Having rejected the identification of truth with what he calls rational acceptability, Putnam there suggests that

truth is an *idealisation* of rational acceptability. We speak as if there were such things as epistemically ideal conditions, and we call a statement “true” if it would be justified under such conditions.¹⁷

He explains that, as he intends the notion, “epistemically ideal conditions” are an idealization in the same way that frictionlessness is: they are conditions that we cannot actually attain, nor—he adds, interestingly—can we “even be absolutely certain that we have come sufficiently close to them.” He is explicit that he is not “trying to give a formal *definition* of truth, but an informal elucidation of the notion.” And he goes on to say that

the two key ideas of the idealisation theory of truth are (i) that truth is independent of justification here and now, but not independent of *all* justification. To claim a statement is true is to claim it could be justified. (ii) Truth is expected to be stable or “convergent.”¹⁸

Putnam has, of course, since officially moved a long way from these ideas.¹⁹ But this is the nearest that he ever came to explicitly endorsing the Peircean conception, and it is clear that his words left considerable latitude for interpretation. In particular, there was no unmistakable suggestion of a key feature of the Peircean proposal: that some *single* set of “epistemically ideal conditions” would be apt for the appraisal of any statement whatever.

Putnam himself subsequently returned to emphasize that point. In the Preface to *Realism with a Human Face* he again endorsed the idea that to claim of any statement that it is true is, roughly, to claim that it could be justified were epistemic conditions good enough.²⁰ And he goes on to allow that “one can express this by saying that a true statement is one that could be justified were epistemic conditions ideal.” But then he proceeds immediately to repudiate the idea

that we can sensibly imagine conditions which are *simultaneously ideal* for the ascertainment of any truth whatsoever, or simultaneously ideal for answering any question whatsoever. I have never thought such a thing, and I was, indeed, so far from ever thinking such a thing that it never occurred to me even to warn against this misunderstanding. . . . I do not by any means *ever* mean to use the notion of an “ideal epistemic situation” in this fantastic (or utopian) Peircean sense.²¹

Rather, the notion of ideal epistemic circumstances stands in need of specialization to the subject matter under consideration:

If I say “there is a chair in my study,” an ideal epistemic situation would be to be in my study, with the lights on or with daylight streaming through the window, with nothing wrong with my eyesight, with an unconfused mind, without having taken drugs or being subjected to hypnosis, and so forth, and to look and see if there is a chair there.

Indeed, we might as well drop the metaphor of idealisation altogether. Rather, “there are *better and worse* epistemic situations *with respect to particular statements*. What I just described is a very good epistemic situation with respect to the statement ‘there is a chair in my study’.”²²

These remarks might invite the following regimentation. Let us, for any proposition *P*, call the following the *Peircean biconditional* for *P*:

P is true if and only if, were *P* appraised under conditions *U*, *P* would be believed,

where *U* are conditions under which thinkers have achieved some informationally comprehensive ideal limit of rational-empirical enquiry. And let us call the following the corresponding *Putnamian biconditional* for *P*:

P is true if and only if, were *P* appraised under topic-specifically sufficiently good conditions, *P* would be believed.

Then we now have two contrasting pragmatist conceptions of truth to consider. And the question is, Do they—either of them—meet the standard set by our proposed minimalism: do they realize the relevant constitutive platitudes?²³

There is an interesting difficulty about an affirmative answer. Putnam imposed what he termed a convergence requirement on his conception of truth—that there be no statement such that both it and its negation are assertible under epistemically ideal (topic-specifically sufficiently good) conditions.²⁴ This is to be distinguished, of course, from any requirement of *completeness*. The requirement of completeness would be that, for each statement, *either* it *or* its negation must be justified under such circumstances. There seems no good reason to impose any such completeness requirement—no particular reason why all questions that are empirical in content should become decidable under Peirce’s or Putnam’s respective ideal conditions. Indeed, to take seriously the indeterminacies postulated by contemporary physical theory is to consider that there is

reason to the contrary. We should expect that a pragmatist would want to suspend the Principle of Bivalence for statements that would find themselves in limbo under epistemically ideal, or topic-specifically sufficiently good, conditions in this way, and ought consequently, one would imagine, to want to suspend it in any case, failing an assurance that no statements are actually in that situation.

So what is the promised difficulty? That there is, apparently, a simple inconsistency within the triad uniting either of our pragmatist biconditionals with the claim that the notion of truth it concerns complies with the minimal platitudes, and the admission that certain statements may remain undecidable under epistemically ideal, or topic-specifically sufficiently good, circumstances, neither they nor their negations being justified. For, as we have seen, the minimal platitudes impose the standard Negation Equivalence:

(NE) It is true that [not- P] iff it is not true that [P].

And to allow that, even under epistemically ideal or topic-specifically sufficiently good circumstances, we might yet be in a state of information that provided warrant neither for P nor for its negation would force us to reject the right-to-left ingredient in (NE) when “true” is interpreted in accordance with either pragmatist biconditional. In other words, it seems that epistemically ideal or topic-specifically sufficiently good circumstances cannot be neutral both on a statement and its negation if the Equivalence Schema is in force over all assertoric contents, if every assertoric content has a negation that is an assertoric content, and if truth is Peircean or Putnamian.

Simple though this train of thought is, it provides, on the face of it, a devastating blow to both pragmatist proposals. Leave on one side the obvious difficulties occasioned by the undecidability of mathematical examples like, say, the generalized continuum hypothesis. Surely, it should not be true a priori even of empirical statements in general that each would be decidable—confirmable or disconfirmable—under epistemically ideal or topic-specifically sufficiently good circumstances. But the relevant minimal platitudes, for their part, presumably hold true a priori. So *if* either pragmatist proposal were a priori correct—as it has to be if it is correct at all—it would have to be a priori that if a statement failed to

be justified under epistemically ideal or topic-specifically sufficiently good circumstances, its negation would be justified instead—just the thing, it seems, that cannot be *a priori*. Invited conclusion: such proposals incorporate mistaken *a priori* claims about the concept of truth, and the properties they present are hence unfitted to serve as realizers of that concept.

Indeed, the point is more general: a simple extension of the argument seems to tell not just against the two tabled pragmatist proposals but against any attempt to represent truth as essentially evidentially constrained. Someone in sympathy with Dummettian antirealism, for instance, may content herself with a one-way Principle of Epistemic Constraint,

(EC) If *P* is true, then evidence is available that it is so.

Yet still be posed an embarrassment by the argument. For if no evidence is available that *P*, then, contraposing on (EC), she ought to allow that it is not the case that *P* is true, whence, by the Negation Equivalence, its negation must count as true. So in the presence even of a one-way epistemic constraint, the unattainability of evidence for a statement is bound, it appears, to confer truth on, and hence, via (EC), to ensure the availability of evidential support for, its negation—contrary to what, someone might very well think, the antirealist could and should admit, namely, that some statements may be such that no evidence bearing upon them is available either way, even under idealized conditions of investigation. (Indeed, how do we explain the semantical antirealist's characteristic refusal to allow the unrestricted validity of the Principle of Bivalence unless it is based on precisely that admission, coupled with the insistence that truth is evidentially constrained?)

What room does such an antirealist have for maneuver here? We can take it that, unless she decides to off-load the notion of truth entirely, there is no denying the Equivalence Schema. Maybe trouble might somehow be found for the move from that to the Negation Equivalence. But the prospects do not look bright.²⁵ What is needed, rather, is a way to reconcile the Negation Equivalence with an insistence that truth is evidentially constrained and the admission that not every issue can be guaranteed to be decidable, even in principle. But is there any scope for such a reconciliation?

Yes, there is. There can be no denying that the Negation Equivalence commits someone who endorses (EC) to allowing (A):

- (A) If no evidence is available for P , then evidence is available for its negation.

And, of course, it's extremely easy to hear this as tantamount to the admission that evidence is in principle available either for affirming P or denying it. But there is a suppressed premise in this turn of thought: the premise (B), an instance of the law of excluded middle:

- (B) Either evidence is available for P or it is not.

Classically, of course, the conditional (A) is an equivalent of the disjunction (C):

- (C) Either evidence is available for P , or evidence is available for its negation.

But the proof of the equivalence depends on the instance of the law of excluded middle, (B). If we may not assume that evidence either is or is not available for an arbitrary statement, then the convertibility of lack of evidence for a particular statement into evidence for its negation, demanded by the Negation Equivalence when truth is evidentially constrained, need not impose (C), and so need not be in contradiction with the a priori unwarrantability of the claim that the scales of (in principle) available evidence must tilt, sooner or later, one way or the other, between each statement and its negation.

This is a substantial result. It teaches us, in effect, that in order to sustain the claim of our two pragmatist proposals—and indeed any broadly Dummettian antirealist proposal—to offer defensible conceptions of truth, the associated package must include revisions of classical logic *of a broadly intuitionistic sort*. For otherwise there is no possibility of modeling the minimal platitudes consistently with a proper recognition that decidability is often not guaranteed even under ideal—Peircean or Putnamian—conditions.

Provided this way of surmounting the difficulty is accepted, our two pragmatist proposals remain in the field as offering two possible ways in which a property satisfying the minimal constraints on truth may be constructed out of assertibility by idealization. Of course, we have only

considered just one problem, so the proposals' claim to succeed in that regard would need more detailed review. But I shelve consideration of that review to turn to another serious and independent form of difficulty confronting each of them, a difficulty that, I contend, should force a pragmatist-inclined philosopher to look for a subtly different kind of conception of truth.

Here is a generalization (and, in one respect, simplification—see note 28) of an objection advanced by Alvin Plantinga specifically against the Peircean proposal.²⁶ Assume any purported account—or indeed any “informal elucidation”—of truth of the form (o):

(o) It is true that $P \leftrightarrow (Q \Boxrightarrow Z(P))$

where Q expresses a general epistemic idealization, $Z()$ is any condition on propositions—for instance, being judged to be true by the ideally rational and informed thinkers whose existence is hypothesized by Q , or cohering with the maximally coherent set of beliefs whose existence is hypothesized by Q , etc.—and “ \Boxrightarrow ” expresses the subjunctive conditional. Since (o) is purportedly a correct elucidation of a concept, it presumably holds as a matter of conceptual necessity. Thus:

(i) Necessarily (It is true that $P \leftrightarrow (Q \Boxrightarrow Z(P))$)

Now suppose that (ii):

(ii) Possibly ($Q \ \& \ \text{Not } Z(Q)$)

Then, by logic and the Equivalence Schema, (iii):

(iii) Possibly (It is true that $Q \ \& \ (Q \ \& \ \text{Not } Z(Q))$)

But (iii) contradicts (i), with “ Q ” taken for “ P ,”²⁷ which therefore entails

(iv) Not possibly ($Q \ \& \ \text{Not } Z(Q)$)

So

(v) Necessarily ($Q \rightarrow Z(Q)$)

A necessarily true conditional ought to be sufficient for the corresponding subjunctive, so:

(vi) $Q \Boxrightarrow Z(Q)$

So, from (i):

(vii) It is true that Q

So by the Equivalence Schema again:

(viii) Q

The upshot is, it seems, that anyone proposing an account of truth of the shape typified by (o) must accept that the idealization Q *already obtains*. Thus the Peircean must accept that conditions are already “epistemically ideal,” (and a coherence theorist must accept that there already is a controlled, comprehensive, and coherent set of beliefs.)²⁸ Obviously, this is unacceptable. And it is not clear how the Peircean can respond.

However, just here is where there may seem additional point to the more modest Putnamian proposal. For the key to the proof above is the license, granted by the Peircean conception of truth in particular, to assume that the conditions that are ideal for the appraisal of the proposition U are the very conditions depicted by that proposition—it is this assumption that sanctions the substitution of “ Q ” for “ P ” in (o). Suppose instead that, with erstwhile Putnam, the pragmatist drops the idea of such a comprehensive set of epistemically ideal conditions and that (o) gives way to a range of Putnamian biconditionals:

(o') It is true that $P \leftrightarrow (Q_P \leftrightarrow Z(P))$

Here Q_P is the hypothesis that conditions are sufficiently good for the appraisal specifically of P . We can advance as before to:

(iii') Possibly (It is true that $Q_P \ \& \ (Q_P \ \& \ \text{Not } Z(Q_P))$)

But nothing harmful need follow unless one of our Putnamian biconditionals is:

It is true that $Q_P \leftrightarrow (Q_P \leftrightarrow Z(Q_P))$

which will be available only if conditions Q_P are topic-specifically sufficiently good not merely for the appraisal of P but also for the appraisal of the proposition Q_P itself, that is, if $Q_P = Q_{Q_P}$. And why should that be so?

Well, but the question should be, Is it certain such an identity is *never* realized? Consider Putnam's own example: a sufficiently good epistemic situation for appraisal specifically of “There is a chair in my study.” That would be, he said, to be in my study, with the lights on or with daylight

streaming through the window, with nothing wrong with my eyesight, with an unconfused mind, without having taken drugs or being subjected to hypnosis, and so forth. But wouldn't these conditions likewise be sufficiently good conditions in which to appraise the claim that I was indeed in my study, with the lights on or with daylight streaming through the window, with nothing wrong with my eyesight, with an unconfused mind, without having taken drugs or being subjected to hypnosis, and so forth? Maybe not—maybe there is some condition whose addition to the list would not improve my epistemic situation with respect to "There is a chair in my study" but without which I would not be best placed to assess the complex proposition just stated. But even if so in the particular example, must that *always* be so? Unless we can see our way to justifying an affirmative answer, there can be no assurance that Plantinga's problem can be resolved by a fallback to Putnamian biconditionals.

In fact it is clear that the most basic problem with the Peircean biconditional cannot be resolved by this fallback. Plantinga made a difficulty by taking Q for P in (o). But suppose instead we take " Q will never obtain," thus obtaining:

Q will never obtain $\leftrightarrow (Q \square \rightarrow Z(Q \text{ will never obtain}))$

Then if the right hand side is interpreted as in the Peircean biconditional, we have a claim to the effect that conditions will always be less than epistemically ideal just in case thinkers who considered the matter under epistemically ideal conditions would suppose so. This is obviously unacceptable. And it is an illustration of a very general point: that no categorical claim P can be a priori (or necessarily) equivalent to a subjunctive conditional of a certain type—roughly, one whose antecedent hypothesizes conditions under which a manifestation, depicted by the consequent, of the status of P takes place—unless it is likewise a priori (or necessary) that the realization of the antecedent of the latter would not impinge on the actual truth value of the categorical claim. More specifically, it cannot be a priori—or necessary—that

It is true that $P \leftrightarrow$ were conditions C to obtain, such and such an indicator M of P 's status would also obtain

unless it is a priori (or necessary) that the obtaining of C would not bring about any change in the actual truth value of P . For suppose that it is

true that P , but that were conditions C to obtain, it would cease to be so: would M then obtain? Yes. For by hypothesis, P is actually true. So the biconditional demands that M would obtain if C did. So not- P would hold alongside conditions C and M . But in that case M would not be an indicator of P 's status in those circumstances after all. In particular, if M consists in the believing that P , suitably placed thinkers, then the effect will be that their beliefs will be *in error* under conditions C —exactly what the pragmatist proposal was meant to exclude.

This point—or anyway the general thought, epitomized in the phrase, “The Conditional Fallacy,” that subjunctive conditional analyses are almost always unstable—is nowadays very familiar from the literature on dispositions and response-dependence.²⁹ What is clear for our present purpose is that it is no less a problem for Putnamian biconditionals than for Peircean ones. That is, unless it is given a priori that the implementation of conditions Q_P would not impinge on the circumstances actually conferring its truth value on P , it cannot be supposed to hold purely in virtue of the concepts involved that

It is true that $P \leftrightarrow (Q_P \square \rightarrow P \text{ would be believed})$

except at the cost of allowing that even under Q_P circumstances, P might be believed when false. And again, this is just to surrender the idea that belief under ideal circumstances is guaranteed to line up with the facts: the cardinal tenet of this kind of pragmatism.³⁰

4 Pragmatism and Pluralism (II): Superassertibility

The ur-thought behind any pragmatist conception of truth is that the notion should be grounded in ordinary human practices of assessment and epistemic values. So some form of idealized assertibility is the most natural concrete interpretation of the idea. But I think that the Peircean and Putnamian conceptions idealized assertibility in the wrong direction. Warranted assertibility is assertibility relative to a state of information. So it can seem as if there is only one direction for a truthlike idealization of assertibility to assume: to wit, we have somehow to idealize the state of information involved, as both the Peircean and Putnamian proposals do in their different ways. But there is another way. Rather than ask

whether a statement would be justified at the limit of ideal empirical investigation or under topic-specifically sufficiently good circumstances, whatever they are, we can ask whether an ordinary carefully controlled investigation, in advance of attaining any mythical or more practical limit, would justify the statement, and whether, once justified, that statement would continue to be so no matter how much further information were accumulated.

More carefully, another property constructible out of assertibility that is both absolute and, so it is plausible to think, may not be lost—Putnam's two desiderata—is the property of being justified by some (in principle accessible) state of information and then *remaining* justified no matter how that state of information might be enlarged upon or improved. Like Peircean truth, the characterization of this property presupposes that we understand what it is for one state of information to enlarge upon or otherwise improve another. But it does not presuppose that we grasp the idea of a limit to such improvement—a state of information that is itself beyond all improvement—or even have any general conception of what it would be for the topic-specific epistemic circumstances to be unimprovable. So this characterization need not confront questions about the intelligibility and coherence of the idea of the Peircean limit, nor need it confront the question of how appraisal under merely topic-specifically sufficiently good conditions can guarantee the stability of a verdict, and thereby the stability of Putnamian truth.

Elsewhere I have called the property just prefigured *superassertibility*.³¹ A statement is superassertible, then, if and only if it is, or can be, warranted and some warrant for it would survive arbitrarily close scrutiny of its pedigree and arbitrarily extensive increments to, or other forms of improvement of, our information.

This admittedly vague characterization makes purely formal use of the notions of “state of information,” “improvement,” and so on. It's natural to wonder how more concrete yet generally applicable accounts of these notions might be given. But I do not think we need to take these issues on. It is enough for our purposes if the notion of superassertibility is *relatively* clear; clear, that is, relative to whatever notion of warranted assertion is in play in the particular discourse with which we may happen to be concerned. Provided, as in all cases that interest us there will be, there

are generally acknowledged standards of proper and improper assertion within the discourse, there must be sense to be attached to the idea of a statement that under certain circumstances meets the standards of proper assertion and then will or would continue to do so unless the considerations that led to its downfall were open to objection in some way. In short, wherever our discourse displays some measure of convergence about what is warrantably assertible, a corresponding notion of *superassertibility* has to be intelligible. This notion may be unclear in various respects, but they will be respects in which the relevant notion of warranted assertibility was already unclear.

So does superassertibility qualify as a potential truth property—does it satisfy the minimal platitudes? The issues here are actually quite subtle. Let's explore some of the twists. We already noted that superassertibility is, plausibly, both absolute and stable.³² It is uncontroversial that it is potentially divergent in extension from assertibility proper. But it merits consideration whether superassertibility and assertibility coincide in normative force. And the question, anyway, is not merely whether superassertibility has these features but whether they issue in the right kind of way from its sustaining the key platitudes.

Let's focus on the Equivalence Schema. Can a supporter of superassertibility argue compellingly for the validity of (E^s):

(E^s) It is superassertible that *P* if and only if *P*.

If he can, then, as briefly noted above, that will arguably settle the matters of Transparency, Correspondence, and Contrast. The commutativity of superassertibility and negation—the analogue of (NE) for superassertibility—will likewise be a consequence.³³

The matter may seem easily resolved, at least to anyone sympathetic to the idea that for a wide class of admissible substitutions for "*P*," it may be that *P* although no evidence is available to that effect. Such a theorist will want to object that (E^s) cannot be valid, since it conflates right across the board the obtaining of a certain kind of high-grade evidence for *P* with the obtaining of the fact. A suitably chosen proposition—Goldbach's conjecture, say—may be undetectably true, and hence not superassertible, and a suitably chosen superassertible proposition—perhaps that we are not brains-in-a-vat—may be undetectably false. Since (E^s) is hostage to

counterexample, so not a priori true, superassertibility has no case to be a truth property.

But the supporter of superassertibility may rejoin that, quite apart from any doubt about the realism on which it depends, there is something unsatisfactory about the *shape* of this objection. Its claim is that there is no assurance that there are no counterexamples to (E^s). But what does it take a counterexample to be? Is it a true proposition that may not truly be claimed to be superassertible? In that case the objection asserts, in effect, that superassertibility potentially lacks, but as a putative truth property ought to be guaranteed to have, the property of generating a valid equivalence when substituted for “?” in the schema (F):

(F) It is true that it is ? that *P* iff. it is true that *P*

However, (F) contains two mentions of a truth property, which, if interpreted as presupposed by the objection, has to be understood as *distinct* from superassertibility. If that doesn’t seem evident, reflect that while—to one in the cast of mind that fuels the objection—it is a possibility that Goldbach’s conjecture be true without it being true that it is superassertible (provable), it certainly isn’t evident that the conjecture might be *superassertible* without it being superassertible that it is. But if there really can be, as minimalism suggests, a *plurality* of truth properties, qualifying as such by satisfying certain general principles, it is only to be expected that an *illusion* of failure may be created by selective interpretations of “true” as it occurs within those principles. It is as if someone were to argue that physical necessity fails to qualify as a genuine notion of necessity on the grounds that it fails to satisfy the principle Necessarily ($A \leftrightarrow B$) \models Necessarily (A) \leftrightarrow Necessarily (B)

and were then to try to back up that contention by selectively interpreting the final occurrence of “Necessarily” in terms of *logical* necessity. If we wish to determine whether there are counterexamples to (E^s), the proper question to put, the friend of superassertibility contends, is not whether superassertibility satisfies (F), but rather whether it satisfies what results when the two tendentious occurrences of “true” are replaced by ones of “?”:

(G) It is ? that it is ? that *P* if and only if it is ? that *P*

The question is, in effect, whether, whenever it is superassertible that P , it is superassertible that it is so, and vice versa.

Can we arbitrate this exchange? What is suspect about the shape of the original objection can be put like this. If any genuine truth property has to validate (i.e., satisfy a priori) the Equivalence Schema, then clearly, *distinct* truth properties can operate over a single discourse (or range of propositions) only if they are a priori coextensive. Plainly, then, no predicate F can express such a property in a discourse in which it is made to function alongside another predicate G that is already assumed both to validate the Equivalence Schema and to be potentially divergent in extension from F .³⁴ The original objection is therefore cogent only to this extent: to show that a discourse is governed by an evidentially unconstrained notion of truth is, for that reason, to show that superassertibility is not a truth property *for that discourse*.³⁵ But no *global* conclusion is licensed. We have to distinguish the questions (i) whether a predicate's content would enable it, under certain conditions, to function as expressing a truth property; (ii) whether, if so, the relevant conditions are met by any particular discourse; and (iii) whether they are met globally. The objection, drawing as it does on a range of examples where it is thought especially plausible that truth is evidentially unconstrained, is properly targeted against the claim of superassertibility on a positive answer to (iii). But in failing to make any distinction among the three questions, it implicitly begs the other two.

There is, however, on the other side, a similar oversimplification in the suggestion that "*the* proper question to put" is, in effect, whether

(G^s) It is superassertible that it is superassertible that P iff. it is superassertible that P

holds a priori. The right perspective, rather, is this. In the presence of the Equivalence Schema, counterexamples to (E^s) are indeed all and only cases where

(F^s) It is true that it is superassertible that P iff. it is true that P

also breaks down. So if (G^s) is valid, then we know that there can be no such counterexamples, and hence that (E^s) is valid, provided, but only provided *no competitor truth-property* operates alongside super-

assertibility—no predicate, that is, that validates the Equivalence Schema but whose coextensiveness with superassertibility is not guaranteed a priori. If there is a competitor in operation, (F^s) may fail when its occurrences of “true” are suitably interpreted, even if (G^s) is valid without restriction on “P.” If there is no competitor, (G^s) and (F^s) stand or fall together. The status of (G^s) is thus highly germane to question (i). If counterexamples to it cannot be excluded a priori, then there will be no general assurance that superassertibility can function as a truth property even when we give it the fullest elbowroom, as it were—even when we make no initial assumption that a competitor is operating over the discourse. On the other hand, if counterexamples to (G^s) can be excluded a priori irrespective of the range of “P,” then we can return a positive answer to question (i), and the answers to questions (ii) and (iii) will then depend on whether and how widely competitor truth properties should be regarded as in operation.

So is (G^s) unrestrictedly valid? We may return a positive answer if it can be shown that to have warrant for P is to have warrant for the claim that P is superassertible, and conversely.³⁶ The latter direction seems unproblematic. If we have reason to regard a statement as superassertible, then we have reason to think that some (in principle accessible) state of information will stably justify the statement, no matter how added to or otherwise improved. And having reason to think that such a state of information exists is plausibly taken to have the same probative force as actually being in the state of information in question. For instance, proving that a (canonical) proof of a particular statement can be constructed is, as far as probative force is concerned, as good as constructing the proof; and there seems no reason why the point should not survive generalization to the general run of cases where we are concerned with defeasible grounds rather than conclusive ones like mathematical proof.

What is less clear is that to have warrant to assert a statement must be to have warrant to regard it as superassertible. Doubtless, warrant to assert P cannot coexist with warrant to *deny* that P is superassertible, since that would be to have warrant to think that the present case for P would be defeated if we pressed matters sufficiently far, and again, that seems as much as to defeat it already. But the question to ask is, rather, whether warrant to assert P can coexist with lack of warrant to regard it

as superassertible—whether one can coherently combine *agnosticism* about *P*'s superassertibility with regarding a present case for asserting it as sufficient.

I'll outline an argument that the mooted combination of attitudes is *not* coherent, that it is precluded by certain quite basic elements in our ordinary conception of what justification for a statement or, equivalently, warrant for a belief involves. The elements involved are three. The first is that epistemic warranty does not have a sell-by date—what I am warranted in believing I remain warranted in believing *sine die* unless I acquire defeating collateral information. The second is that in warrantably believing any statement *P*, a subject is thereby warranted in believing that a sound investigation, to whatever extent one is possible, would bear her out. The third I shall introduce in a moment.

Suppose I warrantably believe that *P*. Now, what counts as warrant to believe a particular statement varies, of course, as a function of time, place, and background information. So what counts as corroboration of *P* for me if I return my attention to the matter in a year's time, say, may comprise very different considerations to those which warrant my present belief. However, by the first of the two assumptions, I will then be warranted, *ceteris paribus*, in believing *P*; and by the second, I will thereby be entitled to expect whatever sound considerations are then available to me to be corroborative just in virtue of the warrant I possessed a year before.

That establishes a conditional: if I am warranted in believing *P* now, then, if I acquire no other relevant information in the meantime, I will be warranted in future in expecting then-available, sound considerations to bear *P* out. But this conditional is something that I may take myself to know now. So whenever I know its antecedent—which, as remarked earlier, I can whenever it is true, since possession of warrant should be a decidable matter—I can know that in any case where I acquire no further relevant information in the interim, certain expectations will be warranted in future. But to know that certain beliefs will be warranted in the future is, only provided one has no present reason to view them as wrong, to be warranted in holding them now. This is the third element in our ordinary conception of justification advertised above: the firm *promise* of justification for what one has no reason to doubt is already

justification. So to be warranted in believing P involves having justification for believing that any subsequent, soundly conducted investigation, prior to which one has acquired no further relevant additional information, will corroborate P .

This is *close* to the desired result but doesn't quite get it. What would suffice to justify the claim that P is superassertible is warrant for the claim that *any* improvement, I^* , of my present state of information, I , will justify P . But what the foregoing establishes is only that if I am warranted in believing P , then I am warranted in claiming that any such I^* *prior to which I have acquired no further relevant additional information* will justify P . So there is a gap. But perhaps we can eliminate it given the third assumption mooted at the end of the preceding paragraph. Say that a later state of information I^* is *first-time P -incremental* on an earlier one I for a given thinker just if prior to possessing I^* , she has no P -relevant information that she did not possess in I . So our result above was that if I am warranted in believing P in I , then I am warranted in thinking that each I^* that is first-time P -incremental on I will likewise warrant P . And now, in order to extend this result to arbitrary improvements I^* of my present state of information I^* , it suffices to reflect that if I^* is not first-time P -incremental on I , then it must be the terminus of a finite chain, $\langle I, I^2, \dots, I^* \rangle$, each element of which is first-time P -incremental on its immediate predecessor. (The point is simply that no matter what P -relevant information I gather between I and I^* , there has to be a first state of information in which I possess each particular item in it.) Reflect then that, by the result of the previous paragraph, in each I^k in which I am warranted in believing P , I will be warranted in believing that I will be warranted in believing P in I^{k+1} . I can know this in I and hence infer that I am warranted in believing that in I^2 , I will be warranted in believing that in I^3 , I will be warranted in believing ... that I will be warranted in believing P in I^* . Application of the third assumption will then let me simplify to "If I am warranted in believing P in I , then I will be warranted in believing P in any improved state of information I^* ."

The contention that (G^s) holds a priori, without restriction on the range of " P ," is thus very much in play, but I leave it to the reader to satisfy herself of the premisses and detail of this argument, which will

bear a more rigorous examination.³⁷ In general, though, it is hard to see how the making of warranted assertions, and the avoidance of unwarranted ones, could have any distinctive point or consequence unless warrant is taken per se to license expectations about the favourable character of subsequent states of information.³⁸

5 Superassertibility as a Model of Truth

We now need to observe, finally, that it is actually not necessary, in order for superassertibility to qualify as a truth property, that it validate the platitudes unconditionally. It will be of no less significance if superassertibility turns out to validate the basic platitudes only subject to certain *additional assumptions* that, consistently with the platitudes, hold a priori for a particular discourse. Such a finding would put us in a position to say that, whether or not the platitudes are analytic of superassertibility when all occurrences of “true” are so interpreted, it is at least a possible *model* of them: it can be shown to have the features they collectively articulate when they are augmented with suppositions on whose status the platitudes themselves are silent.

How does the inquiry fare if we let it take this direction? One way of pursuing the matter begins by asking what is the relation between superassertibility and knowledge. It would be a tall order to argue unrestrictedly that whatever is superassertible can be known, not merely because one would have to vanquish the metaphysical-realist notion that even an empirically unimprovable theory might simply be mistaken, but perhaps more seriously, because the superassertibility of a statement carries no implication about the *strength* of the available evidence, which, though positive, may be enduringly weak. By contrast, it seems to me a highly intuitive claim that anything we can know is superassertible. Admittedly, this will not be so on any reliabilist conception of knowledge sufficiently extreme to abrogate all connection between knowledge and the possession of reason to believe. On such a view, one can know that *P* just by being a dispositionally reliable litmus of whether or not *P*, even if one has nothing whatever to say in support or explanation of one’s believing or disbelieving *P*. But on any view according to which knowledge requires at least some backup with reasons, that is, with asserti-

bility, it is surely going to require superassertibility too. I do not deny that in suitable circumstances an agent may know something on the basis of information that can in fact be defeated. But if his knowledge claim is not to be undermined by the availability of such defeating information, it is surely required that the negative effect of that information, once acquired, could itself be stably overturned.

Doubtless, the matter needs more discussion. But let me propose (K) as analytic of the concepts of knowledge and superassertibility:

(K) P is knowable $\rightarrow P$ is superassertible

And now suppose we are dealing with a discourse in which, as we conceive, it is guaranteed a priori that each statable truth can, in favorable circumstances, be recognized as such—a discourse for which we can make nothing of the idea that truth might lie beyond all possibility of acknowledgement. Comic and, on a wide class of views about it, moral discourse are each, for instance, in this situation: there seems no sense to be attached to the idea that the comedy of a situation might elude the appreciation even of the most fortunately situated judge, or that the moral significance of an act might lie beyond human recognition, even in principle.³⁹ In any case, suppose that, for each assertoric content, P , in some germane class, we have it a priori that:

(L) $P \leftrightarrow P$ is knowable

Had we the converse of (K),

P is superassertible $\rightarrow P$ is knowable,

the validity for the discourse concerned of the Equivalence Schema for superassertibility,

(E^s) It is superassertible that P if and only if P ,

would, of course, be immediate. But we can skin the cat without appeal to the converse of (K) provided we are entitled to assume one half, as it were, of the commutativity of superassertibility and negation, specifically the direction from

It is superassertible that [not P]

to

Not [P is superassertible].⁴⁰

This principle is equivalent to the inconsistency of the supposition that P and its negation might both be superassertible, and is therefore uncontroversial so long as any two states of information are conceived as mutually accessible and warrant is so conceived that no state of information can warrant contradictory claims.

With this lemma in place, it is easy to see that (E^s) is good. What needs to be shown is that

P , and P is not superassertible.

and

P is superassertible, and Not- P .

are contradictory, just as are “ P and P is not true”, and “ P is true and not- P ”. For the first, merely reflect that if P then, by (L), P is knowable; and if P is knowable, then, by (K), P is superassertible. For the second, reflect that, by the same moves, if Not- P , then Not- P is superassertible, and hence by the commutativity lemma, that it’s not the case that P is superassertible, contradicting the first conjunct. Thus, granted the a priori link between knowability and superassertibility postulated by (K), it follows, for any set of contents that sustain (L) a priori, that the assertion of any of these contents is a commitment to its superassertibility and the assertion of its superassertibility is a commitment to (rejecting any denial of) the content.

Plausibly, then, for discourses all of whose contents are in that case, superassertibility satisfies the Equivalence Schema and, in the light of earlier considerations, thus plausibly presents a model of the basic platitudes.⁴¹ And if what I said about the essential appreciability of the moral and the comic is correct, a presumption is established that moral and comic truth can be taken as species of superassertibility.⁴²

One interesting effect is the perspective in which the semantical anti-realism is now placed that generalizes Michael Dummett’s interpretation of mathematical intuitionism. Dummett’s antirealist, inspired by considerations concerning the acquisition and manifestation of understanding, contends that if the meaning of a statement is to be regarded as determined by its truth conditions, then truth cannot outrun our ability (in principle) to know. But then the thesis is that assumption (L), the equiv-

alence of “*P*” and “*P* is knowable,” holds globally for *all* intelligible assertoric contents. So, granted (K), the semantical antirealist contention becomes, in effect, that truth behaves, or ought to behave *everywhere* in a fashion that allows it to be construed as superassertibility. And to respond to the manifestation and acquisition arguments will be to explain how the currency of a notion of truth that cannot be modeled in terms of superassertibility is distinctively displayed in certain aspects of our linguistic practice, and how such a conception of truth might be arrived at in the first place.

This seems to me a helpful perspective on the Dummettian debate. Semantical antirealism now distances itself from the almost certainly doomed project of attempting a meaning-theory that proceeds in terms of an indexical notion of assertibility. Instead, it avails itself of a notion of truth, contrasting with assertibility, and an associated truth-conditional conception of meaning. But it can do this only because superassertibility is, as any antirealistically acceptable notion of truth must be, an essentially epistemically constrained notion—for if *P* is superassertible, it must be possible to alight on the (de facto) indefeasible state of information that makes it so, and then to accumulate inductive grounds for identifying it as such.

For the purposes of pragmatism, for its part, the crucial reflection is that superassertibility is, in a clear sense, an *internal* property of the statements of a discourse—a projection, merely, of the standards, whatever they are, that actually inform belief formation and assertion within the discourse. It supplies no external norm—in a way that truth is classically supposed to do—against which our ordinary standards might themselves be measured *sub specie Dei* and might rate as adequate or inadequate. Rather, the way in which it is fashioned from our actual practices of assessment renders superassertibility as well equipped to express the aspiration for a developed pragmatist conception of truth as any other candidate known to me. If it seems to distort our thinking about truth in particular regions of discourse to conceive it in such terms, that, it seems to me, will be a measure of the local unnaturalness of pragmatism itself.

Notes

This paper, at Michael Lynch's suggestion, revisits some of the arguments and themes of chapters 1 and 2 of my *Truth and Objectivity* (Cambridge: Harvard University Press, 1993.) I am grateful to Michael Lynch for giving me the opportunity to present these ideas to the readership of the present volume and for helpful suggestions about what best to include. Sven Rosenkranz also gave me extremely detailed and helpful suggestions about both contents and structure. My thanks to Harvard University Press for permission to include excerpts from *Truth and Objectivity* and to the University of Calgary Press for permission to include passages from my more recent paper "Truth: A Traditional Debate Reviewed" (*Canadian Journal of Philosophy*, suppl. vol. 24 [1998]: 31–74; reprinted in *Truth*, ed. Simon Blackburn and Keith Simmons [Oxford University Press, 1999]; the official dates notwithstanding, first published in German in Matthias Vogel and Lutz Wingert, eds., *Unsere Welt gegeben oder gemacht? Menschliches Erkennen zwischen Entdeckung und Konstruktion* [Frankfurt am Main: Suhrkamp, 1999]). The present paper was completed during my tenure of a Leverhulme Personal Research Professorship. I gratefully acknowledge the support of the Leverhulme Trust.

1. That is, the Equivalence Schema and the Disquotational Schema yield instances whose truth is knowable a priori by anyone who is in a position to understand them. As is familiar, the right-to-left directions of these equivalences become contestable if truth-value gaps or many truth values are admitted. This complication is pursued in discussion note 1 of chapter 2 of *Truth and Objectivity*. But I do not think that any deflationist should go out of her way to accommodate it, since rejection of the right-to-left direction of the Equivalence Schema flies in the face of what would seem to be an absolutely basic and constitutive characteristic of the notion of truth, that P and "It is true that P " are, as it were, attitudinally equivalent: that any attitude to the proposition that P —belief, hope, doubt, desire, fear, etc.—is tantamount to the same attitude to its truth. For if that's accepted, and if it is granted that any reservation about a conditional has to involve the taking of *differential* attitudes to its antecedent and consequent, then there simply can be no coherent reservation about $P \rightarrow$ it is true that P .

2. It's an unhappy situation that the leading contemporary theorist of deflationism, Paul Horwich, uses both "minimalism" and "deflationism" to characterize his view. However, both his use of "minimalism" and my contrasting one are now entrenched. Probably nobody is confused.

3. This view is contrary to what is suggested by Horwich, *Truth*, 2nd edition (Oxford: Clarendon Press, 1998), pp. 143–144. Horwich there seems to conflate the substantiality of a property with the feasibility of what he calls a "theory of constitution" for this property, i.e., a theory that identifies this property by means of a noncircular equation of the form " x is true iff x is F ," where " F " is replaced by a predicate that does not contain any semantic terms, a fortiori no cognates of "is true." But that just seems to be a prejudice. It is evident from the example of scientific-theoretical predicates, for instance, that there can be no compelling reason to tie expression of a substantial property to explicit definability.

4. There are deflationists who go so far as to deny that "is true" is a genuine predicate at all, but most deflationists are ready to concede that there is such a thing as the concept of truth. A deflationist proposal of the first kind can be found in Grover et al., "A Prosentential Theory of Truth," *Philosophical Studies* 27 (1975): 73–125.

5. That is, the property of having "true" correctly predicable of them. This is presumably what Horwich has in mind when he says that truth denotes a property in the sense in which "every term that functions logically as a predicate stands for a property" (*Truth*, 2nd edition, pp. 141–142).

6. See Hilary Putnam, "On Properties," in his *Mathematics, Matter, and Method* (Cambridge: Cambridge University Press, 1975), pp. 305–322.

7. Thus, the minimalist opposes Horwich's suggestion that truth presents a special case in that an account of the property (or properties) denoted just coincides with an account of the concept that does the denoting. See his *Truth*, 2nd edition, p. 136.

8. On Horwich's interpretation of "substantive property," such reducibility is precisely a necessary condition for a property to be substantive. His suggestion that minimalism (in my sense) is based on the idea that truth is substantive on this understanding thus misconceives the position. See his *Truth*, 2nd edition, pp. 142–143.

9. Proof: derive the two biconditionals one gets from (ES) by respectively negating both its halves and taking "not-*P*" for "*P*." Transitivity of the biconditional then yields (NE).

10. If they were necessarily coextensive, the Negation Equivalence would have to hold for both if for either.

To offset misunderstanding, two points merit emphasis. First, warranted assertibility is here understood to be a notion that is always relativized to a particular state of information. If no such state of information is explicitly mentioned, claims involving this notion will always be understood to relate to the present state of information. Second, the modality involved in "warranted assertibility" does not signify the potential possession of warrants for an assertion, but the actual possession of warrants for a potential assertion. So in particular, merely provable mathematical statements, for which we so far have no proof, do not qualify as warrantably assertible. I believe a confusion of this distinction drives the criticisms in Neil Tennant's paper "On Negation, Truth, and Warranted Assertibility," *Analysis* 55 (1995): 98–104.

11. Thus, for instance Horwich, *Truth*, 2nd edition, pp. 20–23, 139–140.

12. This is, of course, by no means the end of the dialectic. A supporter of the project of Robert Brandom's compendious *Making It Explicit* (Cambridge: Harvard University Press, 1994) will believe that a suitable account of assertoric content—one sustaining the contrast between the proposition that *P* and the proposition that that proposition is assertible—can be constructed out of truth-free materials, as it were. And in his recent book *Meaning* (Oxford: Clarendon

Press, 1998), Horwich himself tries—as he must—to develop a general account of meaning in which truth plays no explanatory part. I cannot pursue the problems with these approaches here. My own view is that the best deflationist response to the inflationary argument is to concede its immediate conclusion but insist that it shows no more than that the concept of truth is indeed of a dimension of (substantial) success and failure, distinct from warrant, for *each particular* proposition, but that there still need be no *single* thing in which, for any two propositions, such success or failure consists. This is indeed one way of taking the “stock-in-trade” response reviewed above. For pursuit of the issue at least some distance beyond this point, see my “Truth: A Traditional Debate Reviewed,” section IV.

13. Which of these forms of opacity goes with the very concept of truth is, of course, contentious, but not that some do.

14. For elaboration of this claim, see my *Truth and Objectivity*, pp. 24–27.

15. Readers familiar with Michael Smith’s work will note a point of contact here with the conception of a *network analysis*, which he derives from Ramsey and Lewis (see in particular chapter 2, section 10, of Smith’s *The Moral problem*, Oxford: Basil Blackwell, 1994). The principal contrast with the approach to truth here canvassed is that a network analysis has to be based on a comprehensive set of platitudes whose conjunction so constrains the target concept that the replacement within those platitudes of all expressions for that concept by a variable and its binding by the description operator results in a definite description that is at the service of an analytically true identity:

F-ness is the property Φ such that $\{\dots\Phi\dots\&\dots\Phi\dots\&\dots\}$

This effectively supplies a reductive analysis of the concept *F*. An analytical theory, by contrast, need not—though it may—subserve the construction of such an analytically true identity.

16. C. S. Peirce, *Collected Papers*, ed. C. Hartshorne and P. Weiss (Cambridge: Harvard University Press, 1935), vol. 8, p. 139. [See chap. 8.—Ed.]

17. We may take it that this is the notion that is now standardly called “assertibility.” Putnam’s grounds for the rejection are two: first, that truth is, plausibly, timeless, whereas warranted assertibility varies as a function of the state of information (“Truth is supposed to be a property of a statement that cannot be lost, whereas justification can be lost”), and second, that assertoric warrant is, whereas truth is not, a matter of *degree*. Recall that I incorporated these points into the platitudes listed above.

18. *Reason, Truth, and History*, p. 56. [See chap. 11.—Ed.]

19. For a review of the contrast between Putnam’s “middle” and most recent views, see my “Truth as Not Wholly Non-epistemic: Putnam’s Peregrinations,” *Journal of Philosophy* 97 (2000): 335–364.

20. Hilary Putnam, *Realism with a Human Face*, ed. James Conant (Cambridge: Harvard University Press, 1990).

21. *Realism with a Human Face*, Preface, p. viii.

22. *Realism with a Human Face*, Preface, P. viii.

23. In making this distinction between the Peircean and sometime Putnamian conceptions, I intend no judgement about whether it is finally stable. As noted, Putnam's intention was that truth, as he informally elucidates it, is, in contrast to warrant, to be a stable property of propositions across time and a property that is absolute, that is not applicable in varying degree. Plainly, this intention can be fulfilled only if to have warrant for a proposition under "epistemically ideal conditions" (however that phrase be interpreted) involves having a case for it that *cannot be defeated* (else we wouldn't have stability) *or improved* (else we wouldn't have absoluteness) *by any further information*. And the only way of ensuring that both points are met would seem to be to require that circumstances count as epistemically ideal (or topic-specifically sufficiently good) with respect to a particular statement just in case *no further information relevant to a verdict on it exists to be had*.

The force of that idea obviously depends on what "relevant" should mean in such a context. In fact, though, it is difficult to see that the term can impose any real restriction at all. For, as is very familiar, warrant is a highly systematic, holistic property of beliefs: the status of a body of information as support for a particular belief turns not simply on the character of the information and the content of the belief but on what beliefs are held as background. A flash of grey glimpsed in the woods may be evidence of the presence of a squirrel if you take yourself to be in New Jersey, say, but of a wood pigeon if you take yourself to be in Scotland. It is no exaggeration to say that any piece of information may, in the context of an appropriate epistemic background, be relevant to any particular belief. How, in consequence, are we to understand the idea of possessing *all* information relevant to a particular proposition? Doesn't it just have to mean possessing *all empirical information*, period? In this way, and notwithstanding his protestations to the contrary, Putnam's intentionally less extreme proposal may seem to slide inevitably toward the Peircean. But I make no assumption about this in what follows.

24. This requirement is superfluous, presumably, since a statement does not count as justified, in any sense that concerns us, unless the case in its favor dominates anything that counts in favor of its negation.

25. For further discussion, see *Truth and Objectivity*, chapter 2, discussion note 1.

26. Alvin Plantinga, "How to Be an Anti-realist," *Proceedings and Addresses of the American Philosophical Association* 56 (1982): 47–70. Plantinga believed he had Putnam in his sights as well, but there are some issues about that, as we will see (though he would be right in any case if the suspicion expressed in note 23 is sound.)

27. Assuming—surely correctly—that a subjunctive conditional, no less than an indicative, is controverted by the actual truth of its antecedent and falsity of its consequent.

28. Plantinga's version of this argument exploits the S4 principle—that what is necessary is necessarily necessary—to derive the conclusion that the idealization Q holds of necessity. But the derivability of Q , unnecessitated, is quite bad enough. A proponent of the “Peircean” conception, or a coherence account of truth, certainly would not intend that the actual obtaining of epistemically ideal conditions, or the actual existence of a maximally coherent belief set, should be consequences of the account. Indeed, these conditions are precisely thought *not* to obtain—hence the counterfactual analysis.

29. A useful explicit discussion is Robert K. Shope's “The Conditional Fallacy in Contemporary Philosophy,” *Journal of Philosophy* 75 (1978): 397–413. The Conditional Fallacy is, of course, a crucial difficulty for certain classical forms of philosophical reductionism—behaviorism and phenomenalism, for instance—but like another absolutely basic structural problem for such views, the holistic interdependencies discussed in Christopher Peacocke's *Holistic Explanation* (Oxford: Clarendon Press, 1978), seems never to have been clearly appreciated during the heyday of debate about them.

30. Both pragmatist conceptions also confront a distinct worry concerning the implicit assumption that epistemically ideal or topic-specifically sufficiently good circumstances are *unique*. Only if so can the proposed conceptions of truth ensure convergence of opinion under such circumstances. But given that the relation *is evidence for* is holistically conditioned by background empirical theory, what a priori obstacle is there to the possibility that *conflicting* sets of beliefs be arrived at under epistemically ideal or topic-specifically sufficiently good conditions as a result of theorists having successfully maintained distinct theoretical backgrounds throughout the information-gathering process—so that an opinion formed about a particular statement can vary as a function of the direction in which, so to say, the idealized circumstances are approached? This thought is amplified, in rather a different context, in chapter 4 of *Truth and Objectivity*.

31. See my *Realism, Meaning, and Truth* (Oxford: Basil Blackwell 1987), pp. 295–302, and *Truth and Objectivity*, *passim*.

32. In fact, it is stable provided the range of the “states of information” quantifier in its definition is stable. That's an assumption that would be questioned by, for instance, an antirealist about the past, or future, who contested whether we should think of the totality of states of affairs as eternal. But, of course, such an antirealist would regard the truth predicate as unstable in any case, so that, in the view of such a theorist, instability stemming from that source would not disqualify superassertibility as a truth predicate. For further reflections on the matter, see *Realism, Meaning, and Truth*, pp. 300–302.

33. Of the remaining platitudes, Embedding is presumably uncontroversial if all assertoric contents sustain it and all are apt to be superassertible. A degree of Opacity is likewise uncontroversial for superassertibility (though what degree of Opacity any truth property has to display is in any case likely to be a vexed question).

34. This is because it cannot be a priori that (P iff P is F) if it is a priori that (P iff P is G) but not a priori that (P is G iff P is F).

35. Since, trivially, if P is superassertible, there has to be evidence for P .

36. For suppose that to have warrant for A is to have warrant for B and vice versa, but for *reductio*, that A is superassertible, while B is not. Let I be a total state of information in virtue of which A is superassertible, i.e., I warrants A and so does any improvement I^* of I . By hypothesis, I also warrants B . Since B is not superassertible, there must therefore be some improvement I^* of I that fails to warrant B . Since any such I^* warrants A , the supposition is contradicted. This shows that coincidence in assertibility conditions suffices for a pair of statements both being superassertible if either is. So if " P " and " P is superassertible" have the same assertibility conditions, (G^s) follows.

37. A beginning is made in discussion note 3 at the end of chapter 2 of *Truth and Objectivity*.

38. For further discussion of this general thought, see chapter 9, note 13, of *Realism, Meaning, and Truth* and the other passages in that book there referred to.

39. I prescind from the complication that the bearers of comic and moral predicates may be spatially or temporally remote. Naturally, modifiers of time and place throw up the same prima facie barriers to the acknowledgeability of comic, or moral truth, broadly conceived, as they pose for discourses in general. A similar point applies, of course, to quantification.

40. The other direction may easily be established by appeal to (K) and (L) as follows:

- (1) Not [P is superassertible] hypothesis
- (2) Not [P is knowable] 1 (by K)
- (3) Not P 2 (by L)
- (4) It is knowable that [not P] 3 (by L, ' $\text{not } P^*/P^*$ ')
- (5) It is superassertible that [not P] 4 (by K)

A different argument for (E^s) is presented in the Appendix to "Truth: A Traditional Debate Reviewed."

41. Such a conclusion could be drawn locally, of course, even if the general validity of (K) is rejected, provided that knowledge entails superassertibility in at least some discourses of which (L) is a priori true.

42. Only a presumption, though. A discourse that meets the conditions described, and so permits superassertibility to model the platitudes characteristic of truth, may yet have other features that impose differences between the two concepts. Getting clear about what such features could be is exactly what is involved in getting clear how realist/antirealist debate is possible after minimalism about truth is accepted on both sides.