TRUTH AS SORT OF EPISTEMIC: PUTNAM'S PEREGRINATIONS*

Two major changes of mind have characterized Hilary Putnam's philosophy: from the early realism to the "internalism" of the late 1970s and 1980s, and then, in the last decade, back to a qualified—"common-sense" or "natural"—realism supposedly innocent of the objectionable features against which internalism had justly reacted. Someone who wants to understand what is essential to the ingredient positions in this progression, and the motives for the moves from one position to the next, will need to shoulder a very broad philosophical agenda. But some of the key issues concern, of course, the concept of truth. The "metaphysical realism" which Putnam attacked in his middle period was associated with a concept of truth which is evidentially utterly unconstrained—a concept which would permit an empirical theory that was ideal by all internal and operational criteria to be false—whereas internalism proposed a notion whereby truth would coincide with some kind of idealization of rational acceptability. So much is well known. What may seem less clear is in what respects, if at all, purely as far as the concept of truth is concerned, metaphysical realism and Putnam's most recent "common-sense" or "natural" realism should differ.

I shall canvass a possible answer to that question by showing how one influential line of criticism, canvassed by Alvin Plantinga,[1] of what many commentators (mistakenly) took to be the Peircean conception of truth defended in Putnam's *Reason, Truth and History*[2] leads naturally to a modified, though still evidentially constrained conception of truth: a conception which not merely has resources to handle Plantinga's and other recent lines of objection to broadly Peircean accounts, but which also permits the type of "recognition transcendence" of truth to which common sense is attracted and to which Putnam gave his recent blessing in the Dewey Lectures.[3] I shall suggest that this concept of truth, though utterly foreign to

* Thanks to Alvin Plantinga, Sven Rosenkranz, and Timothy Williamson for critical comments.

[1] "How to Be an Anti-Realist," *Proceedings and Addresses of the American Philosophical Association*, LVI (1982): 47-70.

[2] New York: Cambridge, 1981.

[3] "Sense, Nonsense, and the Senses: An Inquiry into the Powers of the Human Mind," this JOURNAL, XCI, 9 (September 1994): 445-517.

0022-362X/00/9706/335–64

metaphysical realism, does properly belong with "the natural realism of the common man" that Putnam now defends (*ibid.*, p. 483).

I

The mistake of the commentators I just referred to may well have been two-fold. For it is unclear whether even C. S. Peirce himself ever actually endorsed exactly that conception of truth which modern commentary thinks of as "Peircean." The passage from Peirce[4] usually cited runs:

> 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 (*ibid.*, p. 139, my italics).

This statement of Peirce's incorporates a number of elements associated by contemporary commentators with the pragmatist tradition—in particular, the repudiation of evidence-transcendent truth, and the implicit inbuilding into the very notion of truth of the cognitive values of human investigators. The *alethic fatalism* of the passage is, however, no part of the "Peircean" conception of truth as it is nowadays most often understood. Peirce seemingly believed in a predestined march toward a stable scientific consensus among "all who investigate"; but the received understanding of the "Peircean" view 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 inquiry to some kind of ideal limit. I do not know whether an unmistakable advocacy of this type of conception of truth—whereby a biconditional like: *P* is true if and only if, were epistemically ideal conditions to obtain, *P* would be believed by anyone who investigated it, is supposed good a priori for all truth-apt claims—is anywhere to be found in the actual writings of Peirce.[5] But the terminology is entrenched, and I shall abide here by the prevailing understanding of what it is for a view of truth to be "Peircean."

[4] *Collected Papers,* Volume VIII, C. Hartshorne and P. Weiss, eds. (Cambridge: Harvard, 1935).

[5] He may have moved to something closer to such a view of truth later in his life. See *Collected Papers,* Volume V, p. 495, where he writes: "Truth's independence of individual opinions is due (so far as there is any "truth") to its being the predestined result to which sufficient inquiry *would* ultimately lead"(my italics).

II

Whatever the truth about Peirce, Putnam should never confidently have been read as a "Peircean." Recall the famous passage in *Reason, Truth and History* which has regularly been so interpreted. Having rejected the identification of truth with what he calls *rational acceptability*,[6] Putnam there suggested that:

> ...truth is an *idealization* 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 (*op. cit.*, p. 55).

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" (*ibid.*). He is explicit that he is not "trying to give a formal *definition* of truth, but an informal elucidation of the notion" (*ibid.*, p. 56). And he goes on to say that:

> ...the two key ideas of the idealization 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" (*ibid.*).

So far as I am aware, this is the nearest that Putnam ever came to explicitly endorsing the Peircean conception, and it is clear that his words leave considerable latitude for interpretation. In particular, there was no clear suggestion of some *single* set of "epistemically ideal conditions," apt for the appraisal of any statement whatever.

Putnam himself subsequently returned to clarify that point. In the Preface to *Realism with a Human Face*,[7] he again endorsed the idea that to claim of any statement that is true—"that is, that it is true in its place, in its context, in its conceptual scheme"—is, roughly, to claim that it could be justified were epistemic conditions good enough. 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 (*ibid.*, p. vii).

But now 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

[6] We may take it that this is the notion which is now standardly called *assertibility*.
[7] James Conant, ed. (Cambridge: Harvard, 1990).

> 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... (*ibid.*, p. viii).

He continues:

> There are some statements which we can only verify by failing to verify other statements. This is so as a matter of logic (for example, if we verify "in the limit of inquiry" that *no one ever will verify or falsify P*, where *P* is any statement which has a truth value, then we cannot decide the truth of *P* itself, even in "the limit of inquiry"...I do not by any means *ever* mean to use the notion of an "ideal epistemic situation" in this fantastic (or utopian) Peircean sense (*ibid.*).

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 (*ibid.*).

Indeed, Putnam now suggests, we might as well drop the metaphor of idealization 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" (*ibid.*).

These remarks might encourage 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 inquiry. 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—the suggestion would be—the view endorsed in the Preface to *Realism with a Human Face*, and thereby offered as a gloss on the *Reason, Truth and History* account, was indeed a biconditional elucidation of 'true'. But rather than endorsing the Peircean elucidation,

Putnam was suggesting that Peircean biconditionals should be dropped in favor of Putnamian ones: we should dispense with the fiction of a single, comprehensive, utopian limit of inquiry and localize the idealization to the particular proposition, *P*, concerned; or better, we should drop the idea of ideal epistemic conditions in any case, supplanting it with that of "very good" conditions for the appraisal of *P*—to be detailed by a constructive account, along the lines illustrated for 'There is a chair in my study', of a topic-specific such set of conditions.

III

Whether or not this is the right account of what was happening in the Preface to *Realism with a Human Face*, it merits remark that the reason Putnam there gave for discarding the Peircean biconditionals is not a convincing one. Let it be true that there are consistent pairs of propositions, *P*, *Q*, such that—whether for broadly logical, or quantum-mechanical, or other reasons—the achievement of knowledge whether or not *P* precludes the achievement of knowledge whether or not *Q*. That, so far as I can see, presents a problem for the idea that the truths and falsehoods are just the propositions that we would know to be true or false under conditions *U* only if we assume that each member of such a pair is *determinate in truth value* in any case—that is, if we assume that the principle of bivalence holds without restriction.[8] But why should any Peircean hold that? Michael Dummett, whose thought about these issues ran parallel in many (not all) ways to Putnam's erstwhile "internalism," has always rightly emphasized that one casualty of the adoption of an epistemically constrained conception of truth is the classical principle of bivalence. Very simply: if truth requires knowability (in principle, at the ideal limit of inquiry, or whatever), then we possess no guarantee that either *P* or its negation is true if we possess no guarantee that either *P* or its negation is so knowable. All statements, then, about which there is no guarantee of any verdict under conditions *U* are statements for which we have no justification for assuming bivalence. And in that case, the fact that we know now that there are pairs of mutually consistent statements not both of which can be known under conditions *U* should provide no motive for thinking that there are (or may be) *truths* that would not be known under conditions *U*.

[8] Putnam realized this, of course: note the occurrence of the words "where *P* is any statement which has a truth value" in the passage quoted above.

IV

Myself, I think it doubtful in any case whether Putnam's view of the late 1980s (the view of the author of the Preface to *Realism with a Human Face*) could be neatly captured merely by emphasizing that it was Putnamian rather than Peircean biconditionals that were to feature in the internalist "elucidation" of truth. If that is what he had meant, he would presumably simply have said so. What we get instead are seemingly deliberately indefinite remarks like:

> I am simply denying that we have in any of these areas [quantum mechanics, moral discourse, common sense material object discourse] a notion of truth that totally *outruns* the possibility of justification.[9]

Later, he offers merely that "the truth and justification of ideas are closely connected" (*ibid.*, p. xi). I think the fact is that Putnam at this time was not satisfied with *any* particular formulation of the evidential constraint on truth which he wanted his internalism to require. The imprecise formulations to which he resorted in this period reflect that dissatisfaction.

The matter may seem academic since, by the early 1990s, Putnam had, by his own admission, ceased to defend any conception of truth in the broadly Peircean tradition.[10] In the Dewey Lectures, indeed, a "recognition-transcendent" conception of truth is repossessed on behalf of the "common-sense" realism which Putnam now defends.[11] Here is an illustrative passage:

> How, then, do we understand "recognition-transcendent" uses of the word 'true', as, for example, when we say that the sentence 'Lizzie Borden killed her parents with an axe' may well be true even though we may never be able to establish for certain that it is?... If we accept it that understanding the sentence 'Lizzie Borden killed her parents with an axe' is not simply a matter of being able to recognize a verification in our own experience—accept it, that is, that we are able to conceive of how things that we cannot verify *were*—then it should not appear as "magical" or "mysterious" that we can understand the claim that that sentence is *true*. What makes it true, if it is, is simply that Lizzie Borden killed her parents with an axe. The recognition transcendence of truth comes, in this case, to no more than the "recognition transcendence" of some killings. And did we ever think that all killers can be recognized as such? Or that the belief that there are certain determinate individuals who are or were killers and who cannot be detected as such by us is a belief in magical powers of the mind (*ibid.*, pp. 510-11)?

[9] *Realism with a Human Face*, Preface, p. ix.

[10] See the Preface to his *Words and Life*, James Conant, ed. (Cambridge: Harvard, 1994).

[11] See especially Lecture III, "The Face of Cognition," pp. 488-517.

These rhetorical questions do indeed evoke a plausibly commonsensical conception of the world and our cognitive situation within it: one according to which our epistemic opportunities and powers are essentially limited in space and time, so that it can be a matter of *sheer good luck* whether evidence is available to us here and now of what took or will take place there and then. But must a conception of truth which is tolerant of this commonsensical conception be one about which Putnam's middle period denial "that we have...a notion of truth that totally *outruns* the possibility of justification" is simply mistaken? I shall argue not: that the spirit of the view about truth which is expressed in the internalist statements Putnam now disowns can accommodate his recent reversion to "common sense." This is where it will help to work through the objection to Peircean accounts advanced by Plantinga in "How to Be an Anti-realist."

V

Here is a generalization (and, in one respect, simplification—see footnote 13 below) of Plantinga's objection. Assume any purported analysis—or "informal elucidation"—of truth of the form:

(o) P is true $\leftrightarrow Q \Box\!\!\rightarrow Z(P)$

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, and so on—and '$\Box\!\!\rightarrow$' expresses the subjunctive conditional. Since this is purportedly a correct analysis—or at least a correct elucidation—of a concept, it presumably holds as a matter of conceptual necessity. Thus:

(i) Necessarily: (P is true $\leftrightarrow$ ($Q \Box\!\!\rightarrow Z(P)$))

Now suppose:

(ii) Possibly (Q & Not $Z(Q)$)

Then, by logic and the equivalence schema, P is true $\leftrightarrow P$, we have that:

(iii) Possibly (Q is true & (Q & Not $Z(Q)$))

But (iii) contradicts (i),[12] with 'Q' taken for 'P', which therefore entails:

[12] 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.

(iv) Not possibly (Q & Not $Z(Q)$)

So:

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

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

(vi) $Q \,\square\!\!\rightarrow Z(Q)$

So, from (i):

(vii) Q is true

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.[13] Obviously, that is unacceptable.

VI

How might a Peircean—or a defender of some other "conditionalized" account of truth—respond?

One germane reflection is that the reasoning after line (vi) depends upon a movement from right to left across (o) and thus would not engage an antirealist who proposed merely a one-way evidential constraint on truth, rather than an analysis or some other allegedly a priori biconditional. But the obvious question is how abstention from the right-to-left direction of (o) might be motivated: Is it after all to be allowed that propositions believed in epistemically ideal circumstances might yet be *false*? In that case, it would seem, an ideal theory could be false—and how could that admission possibly be reconciled with anything in keeping with the spirit of pragmatism?

[13] Plantinga's version of this argument exploits the S4 principle—that what is necessary is necessarily necessary—to drive 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 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.

A prima facie more promising thought is that the problem should not afflict a proponent of Putnamian rather than Peircean biconditionals. For the key to the proof is the license, implicit in the Peircean conception of truth, to assume that the conditions that are ideal for the appraisal of the proposition Q are the very conditions depicted by that proposition—it is that assumption that sanctions the substitution of 'Q' for 'P' in (o). Suppose instead that, with Putnam, the antirealist drops the idea of such a comprehensive set of epistemically ideal conditions and that (o) gives way to a range of Putnamian biconditionals:

(o)′ P is true $\leftrightarrow Q_P \mathrel{\Box\!\!\rightarrow} Z(P)$

where Q_P is the hypothesis that conditions are epistemically ideal—or sufficiently good—for the appraisal of P. We can advance as before to:

(iii)′ Possibly (Q_P is true & (Q_P & Not $Z(Q_P)$))

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

$$Q_P \text{ is true} \leftrightarrow Q_P \mathrel{\Box\!\!\rightarrow} Z(Q_P)$$

which will be available only if conditions Q_P are ideal (or 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?

Fine. But the question should be: Is it certain such an identity is never realized? Consider Putnam's own example: an ideal—or sufficiently good—epistemic situation for appraisal 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 would those conditions not likewise be ideal 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 would not improve my epistemic situation with respect to 'There is a chair in my study' but would *significantly improve it with respect to the complex proposition just* stated. Let the reader try to think of one. But even if there is such a

condition 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 that 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 $\leftrightarrow$ $Q \mathrel{\Box\!\!\rightarrow} Z(Q$ 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. That 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 hypothesises 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 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:

P is true $\leftrightarrow$ 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 P is true, 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. In particular, if M consists in the believing that P by suitably placed thinkers, then the effect will be that their beliefs will be *in error* under conditions C—exactly what the internalist 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.[14] What is clear for our present purpose is that it is no less a problem for Putnamian biconditionals than for Peircean ones. Unless, that is, 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:

P is true $\leftrightarrow$ $Q_P \Box\!\!\rightarrow P$ would be believed

except at the cost of allowing that even under Q_P circumstances P might be believed when false. And, again, that is just to surrender the idea that belief under (topic-specifically) ideal—or sufficiently good—circumstances is guaranteed to track the truth—the cardinal tenet of Putnam's internalism.

VII

To be persuaded, however, by this style of point that *any* broadly internalist conception of truth is unsustainable is, I believe, to overreact. Reflect that Plantinga's original reductio does not need a proposed analysis or "informal elucidation" of truth to get its teeth into but will engage any putatively necessary equivalent of 'P is true' of the appropriate subjunctive form. Even a realist might accept, albeit not in the spirit of any kind of analysis or elucidation of 'true', that there can be a necessary biconditional link between 'P is true' and a counterfactual about the beliefs of a *sufficiently* idealized subject—if, say, to go to extremes, the idealization would ensure that the thinker in question would track *all* truth. But such a realist is, on the face of it, put in difficulties, too. Is not such a biconditional, indeed, characteristic of the ordinary conception of an omniscient God? The intent of the Peircean conception is, after all, precisely that (a dispositional form of) omniscience would be the reward for reaching the limit of rational-empirical inquiry. Even if we doubt that we have any clear conception of what such a limit might consist in, it may nevertheless seem quite clear that a (dispositionally) omniscient being would stand to the truths as occupants of that putative limit would stand if the Peircean conception were intelligible and correct. So Plantinga's argument, if good, lends itself to a proof of the existence of the Christian God, or at least of a being possessing something like His traditional epistemic powers! Clearly, there is something the matter with such an argument.

Let us go carefully. The suggestion is to take as an instance of (o) the result of letting Q be: 'There is a Unique Omniscient Being'—

[14] A useful early discussion is Robert K. Shope's "The Conditional Fallacy in Contemporary Philosophy," this JOURNAL, LXXV, 8 (August 1978): 397-413.

one who believes all truths and no falsehoods—and letting $Z(...)$ be: '...would be believed by the Unique Omniscient Being':

(Omn. 1) P is true $\leftrightarrow$ were there to be a Unique Omniscient Being, P would be believed by the Unique Omniscient Being.

The reasoning sketched then takes us to a proof of the existence of such a being. Plantinga himself might welcome that finding; but not, surely, by this route. So where is the mistake? Are not the true propositions, after all, exactly those which would be believed by such a being?

Well, obviously enough, this argument, too, is running foul of a version of the Conditional Fallacy. Again, no subjunctive conditional with an "indicator" consequent can be strictly equivalent to a categorical proposition if the realization of its antecedent cannot be guaranteed not to impinge on the truth value of that proposition. So an instance of (Omn. 1) has a chance of being a priori correct only if P and 'There is a Unique Omniscient Being' are independent. If there were an omniscient being, He would indeed believe exactly the truths. But it cannot be correct to represent the purport of that by something of the form of (o) if we want both 'There is a Unique Omniscient Being' and its negation to be admissible substituends for P. If the range of P is to be unrestricted, then the claim must be, rather, that a Unique Omniscient Being would believe all and only the truths *that would obtain if such a being existed*: that is the conceptual truth about omniscience which (Omn. 1) is trying—ineffectively—to express. Not (Omn. 1) but:

(Omn. 2) Necessarily: were there to be a Unique Omniscient Being, then (P would be true $\leftrightarrow$ P would be believed by the Unique Omniscient Being)

is what is wanted.[15]

The obvious next thought is: Should not a supporter of the Peircean conception have made the *same adjustment from the start*? The "informal elucidation"—of course, it *cannot* now be an analysis, since no straightforward equivalence is being proposed—should have taken the form not of:

[15] This type of adjustment will be very familiar to a reader who is *au courant* with the move from "Basic Equations" to "Provisional Biconditionals" in the setting of recent discussions of the "Euthyphro Contrast"; see, for instance, my *Truth and Objectivity* (Cambridge: Harvard, 1992), pp. 117-20. The prototype of the move is made, of course, by Rudolf Carnap when he contrasts the definition of a term like 'soluble' with its *reduction sentence*—"Über die Einheitssprache der Wissenschaft," *Actes du Congrès International de Philosophie Scientifique*, Fasc II (Paris, 1936); also his "Testability and Meaning," *Philosophy of Science*, III (1936): 419-71, at pp. 440-44.

(o) P is true $\leftrightarrow$ $Q \mathrel{\Box\!\!\rightarrow} Z(P)$

but of:

(o*) $Q \rightarrow (P$ is true $\leftrightarrow Z(P))$

Of course, we may still want to allow that there are good reasons for localizing Q to topic-specific conditions—for replacing it by Q_P, as it were. But that is fine tuning: the crucial point is that, if the analogy with omniscience—the idea that we would stand to the (topic-specific) truths under (topic-specific) epistemically ideal conditions as an Omniscient Being would stand to the truths tout court—is apt, then the Peircean conception, as usually formulated, involves a simple structural faux pas, whose correction should be entirely consonant with its spirit; and those criticisms of it which cluster around the Conditional Fallacy, rather than being fundamental, do no more than highlight that avoidable mistake.

VIII

The suggestion, then, is that an internalist conception of truth might find its best—most resilient—expression not in the endorsement of Peircean or Putnamian biconditionals, nor in the vague idea that truth is somehow intimately connected with justification—that it cannot "totally outrun" it—but in the specific proposal that *Provisional biconditionals* of the form of (o*) hold for all statements—or at least, for all statements for which internal realism is proposed.[16] The proposal of the *moderate internalist* about a given discourse, that is, is that for each statement P of that discourse, the concept of truth is constrained by a principle of the form:

> Were P to be appraised under (topic-specific) sufficiently good epistemic conditions, P would be true if and only if P would be believed.

And the *global* moderate internalist will hold that *all* truth-apt claims P are subject to an instance of that schema which holds good.

Obviously, more will need to be said. Such a principle is no real constraint on the notion of truth, unless certain controls are imposed on the way in which its antecedent is to be understood. If "sufficiently good" epistemic conditions are understood simply as conditions that incorporate whatever features it might take to enable a thinker to track the relevant facts—if there were no obligation for a constructive

[16] Putnam himself never wrote as though internalism would be anything other than a global position. But that is a matter of the detail of its best motivation—and it is not to be ruled out in advance that the best motives might be selective as between one region of discourse and another.

specification of what conditions those might actually be—then no realism, however extreme, need hesitate to accept Provisional biconditionals everywhere. But such a caveat already applied, of course, to the Peircean and Putnamian biconditionals: there, too, there could be no cause to object—other than on grounds of Conditional Fallacy—to the formulations that result if "ideal" or "sufficiently good" are interpreted in nonconstructive, "whatever it takes" terms. Of course, it is another question exactly what it is, for any particular proposition *P*, to specify *constructively* what it is for conditions to be epistemically ideal, or sufficiently good, for *P*'s appraisal. But the central contention of the moderate internalism we are now considering must be that it holds a priori of each truth-apt proposition—or at least each such proposition of a targeted discourse—that a constructive specification is possible of what sufficiently good epistemic conditions for its appraisal would consist in of such a kind that a suitable Provisional biconditional will hold when so understood; that such a specification will proceed by elaboration of *our ordinary practical understanding* of what would constitute best or good enough conditions for the appraisal of such a proposition; and that the availability of such a Provisional biconditional belongs to the very concept of truth.[17]

[17] To avoid misunderstanding, I am not claiming that the moderate internalist should insist that it has to be possible for each truth-apt claim *P* to provide a detailed and constructive specification C_P such that a provisional equation:

> Were *P* to be appraised under conditions C_P, *P* would be true if and only if *P* would be believed.

will hold a priori. What, for the internalist, is a priori is rather that there *are* constructively specifiable conditions in terms of which such a provisional equation will hold; that is, it is to be a priori that:

> There are conditions which allow of fully constructive specification such that were *P* to be appraised under such conditions, *P* would be true if and only if *P* would be believed.

where a specification counts as relevantly constructive only if it avoids explicit (or implicit) appeal to "whatever it takes" to be right about *P*, restricting itself to detailing and building on our ordinary conception of what it does take, but allowing for empirical additions to that conception. (For it should not be inconsistent with internalism to allow that empirical science may teach us of new sources of potential unreliability of observations, for example.)

All this is vague, to be sure. A fully satisfying account would take steps to remedy that. But the vagueness does not matter for our present purposes: first, because the issues here already arose for Putnam's middle-period position—they are not distinctive problems for *moderate* internalism; and second, because—and this will be the crucial point for the connection I want to urge later with Putnam's most recent views—the vagueness involved is perfectly matched by—indeed, just is—that of the *direct-realist* conception of what it is for something to be visible, tangible, or audible: in short, of what it is for a state of affairs to be directly available to our awareness.

IX

So far, we have merely seen that the moderate-internalist conception of truth is not vulnerable to one very basic type of difficulty in Peircean proposals. The next question is whether the Putnam of *Reason, Truth and History* might as well have proposed this view: Is moderate internalism in any way inferior to Putnam's internalism, in its actual historical form, as a reaction against metaphysical realism? Does it contain the resources to provide a stable and adequate vehicle for Putnam's original critique?

Indeed it does. Metaphysical realism, as the view emerges from Putnam's various characterizations of it during his internalist period, is a complex doctrine, involving a number of separable strands: for instance, that "the world consists of a fixed totality of mind-independent objects," that "there is exactly one true description of the way the world is," and that "truth involves some sort of correspondence between words or thought signs and external things and sets of things."[18] Putnam regarded (and still regards) these ideas as unintelligible. But one awkwardness with the stated characterizations is, to the contrary, that they tend to allow of more or less deflationary or commonsensical readings, by the light of which any reaction against them carries an implausible air of paradox—the last thing that the sensible, less extravagant form of realism for which Putnam was seeking ought to do. Putnam was always clear, however, that, on the intended nondeflationary understanding, such ideas had one quite concrete and discussible consequence: that even an ideal theory (a theory that is "*epistemically* ideal for humans"—ideal by the lights of the operational and internal criteria by which we assess the merit of theories) may nevertheless be, in reality, false.[19] According to metaphysical realism, the success of our cognitive endeavors is a matter of *sheer contingency* in the sense that a proposition's compliance with our very best standards of acceptability stands in no necessary—and, in particular, in no rationally scrutable—relation to its being true. An ideal empirical theory—one unimprovable by reference to any standard on whose basis a change in a theory might ordinarily be supported—may, on such a conception, just be a tissue of falsehoods.

That notion—the sheer contingency of cognitive success—is the target, for instance, of Putnam's model-theoretic argument and is

[18] *Reason, Truth and History*, p. 49. (The metaphysical-realist stance is there characterized as *externalism*.) Cf. "A Defense of Internal Realism," in *Realism with a Human Face*, pp. 30-42, at p. 30.

[19] See pp. 12-13 of "Models and Reality," Presidential Address to the Association of Symbolic Logic 1977, reprinted in Putnam, *Realism and Reason: Philosophical Papers, Volume 3* (New York: Cambridge, 1983), pp. 1-25.

implicitly at stake in his famous discussion of the spectre of brains-in-a-vat. The conception of truth enshrined in both the Peircean and Putnamian biconditionals is a direct repudiation of it. Whatever "epistemically ideal" or "sufficiently good" conditions for the appraisal of an entire theory would be, they would have to be conditions under which the merits of the theory would be rationally evident, so that a theory which met *all* operational and internal criteria for excellence would be bound to be endorsed under such conditions. So for each ingredient statement, *P*, in the theory, the right-hand side of either the Peircean or the Putnamian biconditional would be met, and the biconditional would then preclude the possibility that *P* be false. But now it is notable, if that is right, that the same would be true if the a priori epistemic constraint on truth proposed was given not by Peircean or Putnamian biconditionals but instead by Provisional biconditionals:

> Were *P* to be appraised under (constructively specified) sufficiently good epistemic conditions, *P* would be true if and only if *P* would be believed.

For such formulations still entail, of each ingredient *P* in an ideal empirical theory, that the opinion about it by one who considers it under the appropriate conditions—conditions under which its merits are rationally evident—cannot be other than correct. So, again, there is no possibility that such a theory could be false if Provisional biconditionals are generally in force.

Moderate internalism thus clashes with metaphysical realism over exactly the clearest, single most important issue in Putnam's original internalist crusade: the possibility of a false but empirically ideal theory. More generally, it takes issue with metaphysical realism concerning the latter's entire background conception of the place of human cognitive subjects in the world and the conditions on the success of our cognitive endeavors. For metaphysical realism, it is, to stress, always a sheerly contingent question whether our best cognitive efforts—best by our standards, that is—put us broadly in touch with the world or leave us potentially victim to massive error. For moderate internalism, by contrast, it belongs to the concept of truth that each truth-evaluable statement *P* is associated with conditions of assessment which (a) allow of a constructive specification involving at most empirically acceptable additions to our ordinary idea of what reliable assessment of *P* would demand, and (b) are such that opinions formed under them are *eo ipso* truthful. The position thus squarely confronts what I take to be the core metaphysical realist thought: a conception of the objectivity and independence of the world so fashioned as to leave the connection between the prosecu-

tion of (by our standards) the very best cognitive procedures and the attainment of the truth an ultimately inscrutable matter.

X

So far, so good. It is therefore striking that truth as conceived by moderate internalism—in contrast with the standard understanding of the internalism of *Reason, Truth and History*—can "outrun" the evidence in what may seem a perfectly intelligible and domesticated fashion. Provisional biconditionals ensure that for each statement, *P*, within their range, the obtaining or lack of obtaining of a state of affairs depicted by *P* is humanly accessible under the relevant epistemically ideal conditions. But there is no reason to suppose that the implementation of such conditions has to be feasible in any arbitrary situation. The best—or a "sufficiently good"—situation for the appraisal of what is stated by 'Lizzie Borden killed her parents with an axe' might be that of a contemporary witness of her action, or of Lizzie herself. There is nothing in the moderate-internalist thesis—that each truth-evaluable claim admits of some constructively specifiable set of conditions under which its truth value could be recognized—that entails that such conditions are accessible by an arbitrary subject from an arbitrary starting point in space and time.

There is a basic point that needs emphasis here. Any proposed form of constraint on the truth of a statement in terms of what would be recognizable by thinkers operating under hypothetical idealized circumstances has to allow that the occupation of such circumstances might necessitate a change in the spatiotemporal circumstances of a thinker—or in other aspects of her "point of view"—of such a kind as to call for some measure of re-expression of the target statement. Consider, to take a trivial instance, 'Somebody is standing behind you', said by you of me. That is surely a statement whose truth ought to be recognizable by me under "sufficiently good" epistemic circumstances—and such circumstances will no doubt involve turning round and looking! But once I have turned, the form of words appropriate to the expression of the statement can no longer be what it was, if only because the preposition 'behind' must disappear. Still, it would seem absurd to deny that I can ever verify your claim, 'Someone is standing behind you', simply on the grounds that putting myself in position to do so would necessitate its re-expression. And if that is agreed to be absurd, then it would follow that in all the biconditional forms of constraint which we have so far considered, '*P*' needs to be taken as ranging over claims—*propositions*—rather than entities identified by indexical sentence-type. If this were not so, what was just proposed about the Lizzie Borden example would indeed be quite

beside the point: so long as the range of '*P*' was understood to be over entities specified as essentially involving particular forms of indexical vocabulary—tenses, adverbs of place, and the like—it would remain questionable whether we have any clear, constructive conception of "sufficiently good" conditions to ensure that an opinion about the example would be true just in case it was arrived at under those conditions. But against that, I am setting the thought that it should no more seem to be a cheat to invoke conditions whose obtaining would require a thinker to use the present tense in order to engage the claim about Lizzie than it should seem a cheat, in trying to show that I know how to appraise the claim, made of me, that 'There is someone standing behind you', to about-face and look and see.[20]

It may be rejoined that Putnam's position on evidence-transcendent truth in the Dewey Lectures is more generous than can be accommodated by this moderate-internalist treatment of examples that essentially turn on distance in space and time. Speaking of the extension of our

[20] These remarks may seem to undercut the entire thrust of the sort of antirealist concern about past-tense discourse that Michael Dummett, in particular, has drawn to our attention. Whether they really do is a large issue which I shall not try to engage here. But I think it fair to say that it has to be an assumption of Dummett's concern that our abilities, here and now, to recognize the truth value of present-tensed claims *have no role to play* in an appropriately practical—manifestable—account of our understanding of past-tense discourse. Against that, one might naturally feel that the theory of understanding should treat grasp of the tenses as a holism: that there can be no accounting for our grasp of the present tense in isolation, just as there is no properly understanding the numerals for even numbers without grasping the numeral system in general. That is not to reject Dummett's *ur*-thought that grasp of the past tense must somehow be exhausted in certain manifestable abilities; but it is to suggest that the abilities in question may be no more localizable, so to speak, than those concerned with tensed discourse as a whole.

For all the attention expended on this issue over the last thirty years, it remains obdurately difficult to be sure what to think. But even when Putnam, in his internalist period, came closest to Dummett's views, I do not think his motivation ever required him to regard the kind of maneuver proposed on behalf of the moderate internalist with respect to the Lizzie Borden example as an evasion. The bottom line, so to speak, is that the emphasis on understanding as a complex of practical abilities had better go along with the recognition that abilities can be possessed in circumstances where their exercise is preempted or frustrated for incidental reasons. Of course, one who would ascribe such an ability nonetheless then owes an account—a testable account—of what kind of change in the circumstances would enable the subject once again to manifest her ability. But if the ability in question is one to do with an indexical sentence, the possibility is thereby opened up that such a change in the circumstances might render some of the indexical components in that sentence inappropriate. So the only candidate for a *practical* ability associated with the original will have to be something which associates that sentence with an appropriate reformulation together with a practical manifestation of an understanding of the result.

conceptual abilities brought about by a vocabulary that includes not just tenses but the quantifiers, he adverts to:

> ...the possibility of formulating conjectures that transcend even "ideal verifiability," such as "there are no intelligent extra-terrestrials." The fact that this conjecture may not be verifiable even "in principle" does not mean that it does not correspond to a reality...[I]dentifying understanding with possession of verification abilities...makes it mysterious that we should find these words intelligible.[21]

Here, it may seem clear, knowledge of a truth value may be beyond not merely those who, for reasons of spatiotemporal distance, or other accidents of circumstance, are not in sufficiently good position to appraise the statement in question. Rather, it may be supposed, that type of statement might be essentially beyond *all* evidence, no matter what the cognitive starting point from which a thinker went at it. If "common-sense" realism is tolerant of allowing that it might nevertheless record the fact of the matter, then it surely leaves even moderate internalism behind with that concession.

I think more care is needed with such examples. We may take it that Putnam's remarks about the particular case do not involve reposing any particular weight on the fact that it is a negative existential quantification over an unsurveyable domain. That point entails merely that it cannot be known on the basis of a search, not that it cannot be known at all. It might yet be known—or anyway perfectly justifiably regarded as true—under circumstances of sufficiently good theoretical information. If there is to be any chance of constructing a true Provisional biconditional for such a case, we have to expect that the relevant "sufficiently good" epistemic circumstances should involve a thorough understanding of the physical conditions for the emergence of intelligent life (of whatever form) and enough understanding of the distribution and diversity of matter in the cosmos to allow the assignment of a well-grounded probability to the supposition that such conditions are replicated elsewhere besides on this Earth. If moderate internalism can accommodate the example as a truth-apt claim, then the view has to be that under such informationally fortunate circumstances, thinkers would take the correct view of the question. The move that *would* necessitate parting company with moderate internalism is accordingly of one of two kinds: *either* it is denied that there is any such thing as the appropriate kind of understanding of the conditions for the emergence of intelligent life and of the likelihood of a happy intersection of the relevant

[21] Dewey Lectures, pp. 503-04.

parameters elsewhere in the universe *or* it is granted that one could have all that information and yet the fact of the matter could just be brutely different from what one then fully justifiably took it to be. The first is tantamount to thinking that there is no ideal—or sufficiently good—theory of this particular subject matter; the second to the view that here even an ideal theory could be false. Either possibility would be inconsistent with the existence of an appropriate true Provisional biconditional.

Putnam's text does not make it clear, however, whether he would want to advance either claim. Certainly, moderate internalism can allow that the truth value of 'There are no intelligent extraterrestrials' may transcend our knowledge if the explanation of its doing so is our inability (even in principle) to *access* the needed, sufficiently good informational state—which is not implausible, since accumulation of the data on which to construct the best theory of such matters would presumably call for observation of spatiotemporal regions that are too far away. But suppose Putnam had something stronger in mind: that it is a brute possibility that there simply not be enough information out there, even "in principle," which would enable us to advance to the "sufficiently good" understanding called for—or, alternatively, that even the achievement of that informational state and the construction of a theoretical understanding that would ground a high degree of rational certainty on the issue might still, through sheer bad luck as it were, point us in the wrong direction. If he intended to endorse either of those possibilities, then the question is simply: What did metaphysical realism hold about the nature of our cognitive relations to the world that common-sense realism has now discarded? I shall return to this.

XI

It would be disingenuous not to correct any impression that a tolerance of knowledge-transcendent truth is something that comes only with the shift to *moderate* internalism. The insufficiently recognized fact is that an internalism that favored Peircean or Putnamian biconditionals had the same resources all along! True, Peirce's own explicit view entailed that all truths will eventually be known, and hence that all are knowable. But that is not an implication carried simply by the unrestricted endorsement, for all truth-apt claims, of either Peircean or Putnamian biconditionals. Those formulations do no more than ensure that opinions formed under "epistemically ideal" or (topic-specific) "sufficiently good" circumstances will track the facts. To derive from them that any truth is *knowable*, two lemmas would need to be established: first, that opinions formed under such circum-

stances would thereby amount to knowledge; and second, that an opinion formed under such circumstances is always a *possibility* (where the modality involved is the same as that involved in the relevant concept of knowability). If one is externalist enough to allow that any true opinion formed under what are, in fact, appropriately fortunate circumstances counts as knowledge, then the first consideration need be no obstacle—though the issues concerning such epistemological externalism are controversial and substantial.[22] We can set them aside, however, since the second lemma is the crux. Let it be that a proposition is true just in case I would know it if I could appraise it under the right circumstances: it remains that it may be impossible to know it precisely because I *cannot* achieve those circumstances. Thus, an internalism based on Putnamian or Peircean biconditionals can and should take just the same view of evidence transcendence as moderate internalism. Whether or not Putnam himself was sufficiently clear about the point, it is evident that his middle-period internalism need never have been antithetical to the idea of recognition-transcendent truth as such. What it was essentially antithetical to is the notion that recognition-transcendent truth can arise for some reason other than the unavailability of sufficiently good epistemic conditions (or—if it matters on one's preferred account of knowledge—the unrecognizability of such conditions when they arise). But in this respect, that form of internalism and moderate internalism are exactly on a par.

XII

Nevertheless, there is a more specific—and much-discussed—problem to do with recognition transcendence on which an internal realism based on Putnamian biconditionals arguably founders, but with which moderate internalism is perfectly comfortable. I am referring to the so-called *paradox of knowability*, due to F. B. Fitch,[23] concerning claims which are very plausibly taken as truth-apt yet entail limitations on their own knowability. Take $\mathcal{F}$ as some proposition of the form: Q and no

[22] Ought knowledge not to demand, in addition, an assurance that the circumstances in question *are* so fortunate and, as Putnam himself earlier emphasized, is there any guarantee that such an assurance would have to be available, that occupants of sufficiently good epistemic circumstances would have to be in position to know that they were? (Of course, there would be potential difficulty with that question if "sufficiently good" epistemic circumstances are Peircean, that is, if one such set of circumstances is supposed to suffice for all cases. For then the idea that we might occupy such circumstances yet not be able to realize it is the idea of a truth that would not be believed under sufficiently good epistemic circumstances. For related discussion, see my *Truth and Objectivity*, p. 46.)

[23] "A Logical Analysis of Some Value Concepts," *Journal of Symbolic Logic*, XXVIII (1963): 135-42.

thinker will ever rationally believe that Q.[24] We would ordinarily suppose that there must be *countless* truths of this kind—for there are countless propositions, including ones we could, if we chose, quite straightforwardly decide, which are true but will never actually be (rationally) believed by anybody, because nobody will ever actually take the trouble to check. For a Putnamian internal realist, a proposition like $\mathcal{F}$ can, in the nature of the case, never be believed under the relevant "sufficiently good" epistemic circumstances.[25] For such circumstances must include the rationality of the thinker. But rational acceptance of $\mathcal{F}$ would require rational acceptance of both its conjuncts, and rational acceptance of either conjunct under "sufficiently good" epistemic circumstances precludes rational acceptance of the other.[26] So whatever specific "sufficiently good" conditions are germane to the right-hand side of a Putnamian biconditional for a particular such $\mathcal{F}$, it would appear that the resulting subjunctive conditional will never hold good. Someone who operated under sufficiently good circumstances never would (rationally) accept $\mathcal{F}$. Putnamian internalism would seem, accordingly, to have no resources to explain the possibility that such an $\mathcal{F}$ could be true in the first place.

For moderate internalism, by contrast, Fitch-type examples bring out no more than that there are some propositions whose very appraisal under "sufficiently good" circumstances would be sufficient to bring about their untruth. For instance, "sufficiently good" circumstances for the appraisal of $\mathcal{F}$ as above would have to be circum-

[24] Fitch actually took: Q and no-one will ever know that Q. But it is more convenient in the present context to work with rational belief. (But doing so necessitates some departure from Fitch's own reasoning, which exploited the factivity of knowledge.)

[25] Do there even have to be such circumstances in this case? As noted, Putnam observed that "sufficiently good" epistemic circumstances for the appraisal of one proposition may exclude "sufficiently good" epistemic circumstances for the appraisal of another. But I see no reason to suppose that the conjuncts of a Fitch example provide such a case. And—the important point—if they did, then there necessarily would be no "sufficiently good" circumstances for the appraisal of the Fitch conjunction, and the relevant Putnamian biconditional would consequently get $\mathcal{F}$'s truth conditions wrong. (For whatever one's preferred verdict about counterfactual conditionals with necessarily false antecedents, it will presumably be uniform; but $\mathcal{F}$ is a contingency.)

[26] To spell that out, suppose C constitute sufficiently good epistemic circumstances for the appraisal of 'Q and no thinker will ever rationally believe that Q', and hence for both its conjuncts. If Q—the first conjunct— is rationally believed under C, then it is not true that no thinker will ever rationally believe that Q, so not rationally believed either—since, remember, the internalist position requires that, if either conjunct is rationally believed under C, then it is true. On the other hand, if under C the second conjunct is rationally believed—namely, that no thinker will ever rationally believe that Q—then indeed no thinker ever will; so, in particular, there will be no rational belief that Q under conditions C.

stances in which one was in position to appraise each conjunct of it. So, on the one hand, one would be in position to carry out the procedure appropriate to determining whether or not Q is true, and, on the other, one would be in a state of ongoing reflective lucidity concerning what one was/was not in position rationally to believe. Maintaining the latter through implementation of the former procedure would thus have one of two effects: either one would consciously verify Q and, coming thereby rationally and self-consciously to believe it, consciously falsify 'No thinker will ever rationally believe that Q' and thereby $\mathcal{F}$; or one would consciously falsify Q and thereby consciously falsify $\mathcal{F}$ directly. Either way: if $\mathcal{F}$ were to be appraised under sufficiently good circumstances, it would not be true. But that result is perfectly consistent with:

> If $\mathcal{F}$ was to be appraised under sufficiently good circumstances, $\mathcal{F}$ would be true just in case it was believed to be true.

precisely because under those circumstances, as noted a moment ago, $\mathcal{F}$ would not be believed either. Thus, in contrast with the situation with Putnamian biconditionals, the Provisional biconditional for $\mathcal{F}$ is indeed an a priori truth. Provisional biconditionals accordingly involve no misrepresentation of the intuitive truth conditions of such examples, which are perfectly compliant with the constraint on truth which moderate internalism involves precisely because both their being true, and their being believed, are inconsistent with their appraisal under optimal conditions.

Under this perspective, the Fitch paradox emerges as, in effect, a special type of Conditional Fallacy difficulty for the Peircean and Putnamian biconditionals: just one more example of the kinds of awkwardness that can be generated for them by nonindependence between the target proposition and the antecedent of the relevant subjunctive conditional—and, like such problems in general, it is finessed by the shift to Provisional biconditionals. It remains, to be sure, that such a proposition can be true only if an appraisal of it under sufficiently good circumstances never takes place, so that its truth is, in that sense, essentially recognition transcendent. But that there are such recognition-transcendent truths—like the case of Lizzie Borden and in contrast with the situation of 'There are no intelligent extraterrestrials' when pushed beyond the reach of even the very best theory—is indeed merely common sense and nothing which a recoil from metaphysical realism should provide any motive to deny.

XIII

Let us take stock. So far I have offered considerations in support of a number of claims:

That a moderate-internalist conception of truth would have served as well as any conditionalized, "Peircean" proposal as a vehicle for Putnam's middle-period opposition to metaphysical realism.

That a moderate-internalist conception of truth can handle certain well-known structural difficulties for Peircean (and other essentially evidentially constrained) conceptions of truth, including Plantinga's problem and Fitch's paradox.

That a moderate-internalist conception of truth is friendly to at least two important classes of potentially recognition-transcendent truths, typified by the Lizzie Borden and Fitch examples.

All of that is to suggest that Putnam's middle position allowed of stronger play than perhaps it historically received; and that it need not have been inhospitable to certain ideas that it has later seemed important to him to accommodate and emphasize. What is still to be clarified is how far, if at all, Putnam's recent views genuinely demand a departure from this refashioned version of those of *Reason, Truth and History* and *Realism with a Human Face.* If moderate internalism had been the view which Putnam had adopted at the time of *Reason, Truth and History,* would there have been any need for the *further* recantation of the early to mid- nineties? Or does moderate internalism have a case to be regarded as faithful to "the natural realism of the common man"?

What precisely are the key differences between the philosophical syndrome which for Putnam constitutes "metaphysical realism" and "the natural realism of the common man"? Earlier, we noted some aspects of Putnam's own conception of the contrast: the idea of the world as consisting in some fixed totality of objects (and hence of some fixed totality of things falling under some uniform conception of "object"), the conception of the variety exhibited in the world as determined by some fixed totality of properties (and hence of any interest relativity in our classifications as a matter simply of *which* such properties we have chosen to take an interest in), the conception of truth as consisting in some fixed relation of correspondence between our sentences, or thoughts, and the world (and hence of ascriptions of truth as making, more than a claim about the world, claims about the obtaining of such a relation of correspondence), the conception of a sharp and meaningful contrast—dictated of course by the metaphysical-realist ontology of genuine properties—between the areas of discourse where we *describe* how matters objectively stand and areas of discourse where we rather "project" aspects of our own reactions onto the world—all these notions Putnam continues, with his common-

sense realist hat on, to reject as unintelligible.[27] These aspects of Putnam's continuing "recoil," however, do not seem directly to help with the specific question: How, if at all, do the conceptions of truth respectively involved in metaphysical realism and common-sense realism differ? Certainly, common-sense realism will not accept the particular version of the correspondence theory of truth held by metaphysical realism, since it will not accept the associated "metaphysical" ontology of objects and properties. But the key question specifically concerns the evidence transcendence of truth: Does a return to "the natural realism of the common man" involve any kind of disagreement with metaphysical realism about that?

I believe that it does. The issue of evidential constraint—of whether truth requires the possibility of verification and, if so, what notion of possibility of verification is germane—has loomed so large in the secondary literature that almost no thought has been given to what range or gradation of views may open up if a negative answer is returned. But admissions that truth may be evidence transcendent may be driven by two quite different kinds of motive. One such motive is benign. It is indeed part of common sense to view ourselves as finite creatures, whose knowledge-gathering powers are limited. We are limited by our range of sensory susceptibilities—our vision and hearing, for instance, are restricted to quite small bands of frequencies of light and sound. We are limited by our intellects: there are surely many problems that we are too slow or unimaginative to solve. We live short lives, and the spatial distances within which we can investigate are further restricted by the slowness of our means of travel. We cannot travel in time at all (other than forwards at the usual rate). All these ideas go to the core of our common-sense self-conception: it thus goes with that conception that there can be no end of possible ways things might be which we can fully understand—since they allow of characterization within our conceptual repertoire—but which we cannot know to obtain or not because we are grounded in the wrong place, or the wrong time, or because the issues raised are too difficult for us. A conception of truth that denies this, like the kind of antirealism about the past canvassed in Dummett's famous papers,[28] will indeed seem paradoxical and revisionary; for instance, the idea that there are no truths about the past save those for which confirming evidence is available in the present just

[27] See especially the first of the Dewey Lectures, *passim*, and in particular footnote 41 at p. 463.

[28] The *locus classicus* for Dummett's treatment of the issues is his "The Reality of the Past," *Proceedings of the Aristotelian Society*, LXIX (1969): 239-58.

flies in the face of the intuitive conception of the past as a *removed present*, a realm of states of affairs no less determinate than the present but no longer directly accessible to us and whose surviving traces are partial and contingent.

This benign form of evidence transcendence is conceived as going with—in the broadest sense—*contingencies of epistemic opportunity*: in all cases where we have a conception of this kind of how the truth value of a particular statement could be unverifiable, a developed specific account of that conception will consist in detailing limitations of opportunity, or spatiotemporal situation, or perceptual or intellectual capacity, which stop us getting at the relevant facts but to which we, or others, might easily not have been subject—or at least, to which we can readily conceive that an intelligible form of investigating intelligence need not be subject. The hallmark of this general range of cases is thus that in providing an account of how and why it is that they may transcend our epistemic powers, we shall simultaneously begin to provide an account of a suitable set of conditions to figure as the antecedent of an appropriate Provisional biconditional: a set of conditions that creatures very much like us, but situated elsewhere in space and time, or finitely idealized in certain respects, could occupy and, by doing so, ensure that a judiciously formed opinion about the relevant matter would be correct.

The opposed, malignantly-motivated cases arise quite differently. We are concerned here not with contingencies of opportunity but with putative *necessities of limitation*. And it is here that we find a connection with perhaps the single most emphatic theme in the Dewey Lectures: the repudiation of any form of "interface" between the knowing, perceiving, or thinking subject and the worldly subject matter of his thought.

Winning through to natural realism—

Putnam writes:

> is seeing the *needlessness* and the *unintelligibility* of a picture that imposes an interface between ourselves and the world.[29]

The "interface" conception is, of course, very old. It is just a generalization and intensification of the predicament of the thinkers in Plato's Cave. On the interface conception, large sweeps of fact are essentially inaccessible not just to us (Cave dwellers) but to *any* sentient, intelligent creature, no matter where and when situated. The classic way for this to

[29] Dewey Lectures, p. 487; see also p. 505.

happen, of course, is by restriction of all sense experience in its very nature to a domain—of sense data, or mere seemings otherwise construed than as sense data—which are then conceived as standing in merely causal, potentially inscrutable relations to an ulterior material reality. Knowledge of the external world is thus made to rest upon a chancy, backward inferential leap across those causal relations. An exactly analogous predicament arises, of course, on a dualist conception of the mental with respect to knowledge of other minds; and also with respect to knowledge of the past, when memory is conceived merely as the tokening of memory images, conceived as phenomenal proxies, in consciousness. In all these cases, evidence transcendence is not down to contingencies of opportunity. Rather, it is held, we have no coherent conception of the workings of perception, or memory, or of the recognition of others' mental states, save in terms which see the proper operation of the faculties in question precisely as stopping short at an *interface*: that is, at a domain of facts which are *categorically distinct* from those in which we are really interested, and which stand to the latter merely as effects to strictly inscrutable causes. On this type of view, evidence transcendence is a product not of contingencies of opportunity, or of contingent limitations of our powers, but of *metaphysical* shortcomings and divides: consciousness—unless it is the consciousness of God, of which we have no satisfactory conception—is necessarily and essentially insular.

The essence of metaphysical realism, we might say, is thus *interface realism*. And the evidence-transcendent conception of truth which metaphysical realism brings in train is of a malignant kind, the kind that goes with an interface conception of mind's interaction with the world. About these ideas, there is nothing "natural" or "commonsensical." Rather, it is supposed to take philosophical sophistication to appreciate their point. (They are naturally viewed as Cartesian. But for Putnam, the structural mistake that they embody is a cardinal feature of much modern philosophy of mind—what he therefore calls "Cartesian materialism"—which thinks of "representations" as brain states whose connections with what they represent are essentially merely causal.)

If it is no part of common-sense realism to accept this picture of cognitive alienation, to tolerate interfaces, then the evidence transcendence that needs to be provided for should be solely of the benign sort. But the benign conception has it that no genre of states of affairs is essentially beyond our powers of knowledge: under the right—in principle constructively specifiable "sufficiently good"—circumstances, our powers will "reach out" to the very facts in question, and opinions formed will be correct. So, the holding of an appropriate range of Provisional biconditionals should belong with our very conception of the

states of affairs in question, and will thus be a priori. Conversely, if interfaces are allowed, then whatever states of affairs are open to our perceptual, or other faculties will constitute *distinct existences* from those behind the interface, and even if any Provisional biconditionals do hold good for the latter, they will do so only courtesy of a harmony with the former for which there will be no a priori ground. In brief, the epistemology of direct realism, surely part and parcel of "the natural realism of the common man," goes with the Provisional biconditional as an a priori constraint upon truth. If we want no truck with interfaces, then every kind of state of affairs of which we are capable of conceiving must be one of which we can attain a concrete conception of what kind of powers and situation would place us, or relevantly similar beings, into a position from which it would be assured that our, or their, opinion that it did, or did not obtain, would be correct. That is exactly what the possibility of *direct* cognition demands.[30]

[30] One line of resistance to the foregoing, put to me by Timothy Williamson, is worth mention. Direct realists take the view that perception is a certain kind of *tracking ability*: we have the ability to be immediately receptive to aspects of our local environment. They also hold, presumably, that the successful exercise of this ability is the norm. Since their thesis is an article of philosophy, it is presumably considered to hold good a priori (even if the materials to counter various forms of skeptical attack are not immediately to hand). But is saying that our senses normally put us in direct touch with our local environment a commitment to holding that, under in-principle constructively specifiable circumstances they *always* do, as would be demanded by the holding of the appropriate kind of Provisional biconditional? Might not episodes of cognitive dissonance beset even the best perceivers among us, so that occasional misfires occur even under the most favorable conditions, perceivers forming false beliefs, or failing to form appropriate true ones, for no reason anyone can produce? Surely, it is no part of the concept of possessing an ability to provide a *guarantee* that, under some specifiable set of suitable circumstances (including dominant motivation of the subject) successful performance is entrained: perhaps even the most adept cyclist would sometimes fall off her bicycle under perfectly unremarkable conditions.

It may be tempting to reply that, if that were so, it would merely be because of our imperfect understanding of the conditions sufficient for success: that it should always in principle be possible to isolate some further specification of the conditions under which such an aberration would occur and thereby, correlatively, of conditions under which a given ability will invariably show in performance. But how could it be a priori that this is so? (What if the most basic laws governing certain relevant brain events are merely statistical?)

The question is pertinent. I think, however, that we are entitled to proceed, as I have, without engaging this complication. For the wider question is not whether direct realism must involve a commitment to there being true, in-principle constructively specifiable suitable Provisional biconditionals but whether its commitments concerning truth and knowability coincide with those of moderate internalism. And it should be clear that the objection poses no obstacle, actually, to the coincidence. Let it be that for indeterministic reasons—if I may so put it—any human ability fails to guarantee successful performance even under the most favorable conditions. Then this is a complication which the entire internalist tradition should have reckoned with from the start, and none of the bicondition-

It follows that it must still be part of the stock-in-trade of common-sense realism, if it is to be a *thoroughgoing* direct realism, to hold that an empirically ideal theory could not be false. For to suppose otherwise, to suppose that a theory which was operationally and internally ideal—which saved all empirical phenomena and was best by all criteria by which we might assess the competing claims of pairs of empirically adequate theories—might nevertheless be false would be exactly to conceive of its proper subject matter as standing behind an interface: it would be to conceive of the states of affairs in virtue of which its distinctive theoretical claims were true or false as items for our knowledge at best only via proxies.

In sum, the crucial difference between the philosophy of truth of common-sense realism and that of metaphysical realism concerns the proper way of conceiving of the *possible sources* of evidence-transcendent truths. For common-sense realism, they arise because of contingencies of epistemic opportunity. For metaphysical realism, they arise because—according to that view—it is in the nature of most types of states of affairs to influence consciousness, if at all, only via proxies. It goes with the latter conception, just as Putnam always said, that a theory of such states of affairs which is ideal by our standards can be false. It goes with the direct realism of the former conception that appropriate provisional biconditionals should hold, and hence that an ideal theory cannot be false. Thus, the latter contention

---

als—Peircean, Putnamian, or Provisional—which we have canvassed ever had any chance of holding a priori and absolutely. Rather, each could only hold in some appropriately *statistical* form. A Peircean should have proposed, for instance, that the true is what thinkers judging under epistemically ideal conditions would *mostly* take to be true. Putnam's internalist should have proposed that the proposition that there is a chair in my study is true if and only if were it appraised under topic-specifically sufficiently good conditions, it would be believed in the (overwhelming) majority of cases. And a moderate internalist about some region of thought should hold, for instance, not that under certain constructively specifiable circumstances, thinkers' judgments whether P and the fact of the matter would be covariant but that they would *usually* be covariant—that a series of trials would converge on one verdict and that cases of conflict would be isolated and unlikely. But just that, presumably, is what a direct realist who takes seriously the mooted complication will hold about our perceptual thought.

I leave it to an enthusiastic reader to consider in detail whether the dialectic of this paper would have differed in any significant respect had it been organized in terms of these various statistical conceptions of evidential constraint. But notice at least that the Conditional Fallacy difficulties besetting the Peircean and Putnamian proposals still apply; for instance, 'Epistemically ideal conditions will never obtain $\leftrightarrow$ were epistemically ideal conditions to obtain, thinkers would mostly judge that epistemically ideal conditions will never obtain' is still a singularity for a would-be Peircean. Thus, the passage to the (correspondingly statistical) Provisional biconditionals proposed by the moderate internalist will still be well motivated.

should still be part of the common-sense realism that Putnam now advocates, if that position is to incorporate a thoroughgoing direct realism. And a moderate-internalist conception of truth should be recognized as part of "the natural realism of the common man." That is my principal conclusion.

XIV

If these ideas go in the right direction, then it would be unhappy to see the thrust of Putnam's most recent views as a partial rapprochement with metaphysical realism after a period of (over)reaction against it. Rather, they are best seen as a continuing working out of that reaction in a consistent direction. And an epistemic conception of truth, of the moderate-internalist sort, should be a continuing part of the picture. But what remains to be settled, of course, is *how big* a part: whether a direct-realist epistemology, and consequent endorsement of a moderate-internalist conception of truth, should indeed be generalized across the whole gamut of possible states of affairs of which we can claim any clear conception. Putnam spoke of the needlessness and unintelligibility of the interface picture. Did he mean, "Everywhere"? It is essentially the interface picture that we will need to make sense of the idea that 'There are no intelligent extraterrestrials' might be true beyond all evidence available in principle, beyond the verdict of even the very best possible empirical theory. Can one have a stable and coherent *mixed* view—one with merely local interfaces, so to speak, restricted to areas of discourse where knowledge, if attainable at all, essentially depends, even in the very best conceivable case, upon defeasible inference? Those like Putnam (and, of course, John McDowell, by whose John Locke Lectures[31] Putnam's recent work is influenced) who have wanted to press the claim of a sophisticated direct realism (the "second naiveté" Putnam approvingly finds in J. L. Austin) still owe a statement, it seems to me, of the point, if any, beyond which they do *not* want to press the direct-realist claim—and indeed an account of how permitting *any* boundary could spare our conception of what lies beyond it the range of problems—to do, for example, with content and reference—which they urge against those philosophies that would confine our *direct* knowings within more traditionally conceived (interior) boundaries.

CRISPIN WRIGHT

University of St. Andrews
Columbia University

[31] Published as *Mind and World* (Cambridge: Harvard, 1994).