LECTURE

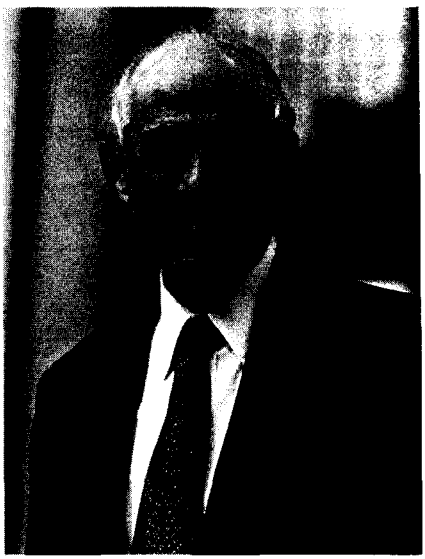
"To Think with Integrity"

Hilary Putnam's Farewell Lecture

This is the last lecture Professor Hilary Putnam gave at Harvard before retiring. He delivered it on May 4, 2000 in Emerson 305 as a conclusion to his legendary course, Philosophy 154: Non-Scientific Knowledge.

The editors would like to thank Anthony Corsentino for his help in preparing this lecture for publication. In order to preserve Professor Putnam's lecturing style, the transcript of the lecture has been edited only slightly.

I REALIZED THIS MORNING THAT I DON'T WANT TO GIVE THIS lecture, and of course it doesn't take much self-knowledge or psychoanalytic penetration to understand that what that really means is that I don't want my career at Harvard to be over. But at this moment, no matter what I may rationally think (and I think I made the right decision), something in me is sad.



However, what's done is done; so I will give a last lecture. I heard somebody give a very elegant acknowledgment lecture just last week beginning with the words, "I am literally speechless." It crossed my mind to say that too, but I won't do that.

The only thing that remains is to continue doing what I've been doing at Harvard for thirty-five years, and at one place or another ever since I taught my first course in 1952 in Evanston: be a pedagogue. So I will do some teaching.

One topic in my Dewey Lectures, Part I of *The Threefold Cord*, that I didn't say much about (and there are many that I didn't get to in this course), is the topic of truth. And I will just say a little about that and then try to say something about the more general themes that have dominated my work since *Reason, Truth, and History.*

In the third Dewey Lecture, I distinguished between two versions of what

Hilary Putnam, widely considered one of the major philosophers of the last fifty years, is Cogan University Professor Emeritus at Harvard. Professor Putnam, whose tremendously influential contributions to philosophy are too numerous to even list here, is the author, most recently, of The Threefold Cord.

today is called 'disquotationalism' or 'deflationism'. One of these I attribute to Frege. The other is, I said, most clearly presented in a little-known article of Rudolf Carnap's, which I refer to in a footnote. It is an article that Carnap published in what was perhaps the most influential anthology in analytic philosophy for many years: the first edition of Feigl and Sellars's *Readings in Philosophical Analysis.* It seems to have been something that Feigl got Carnap to put together from, perhaps, two different pieces that had appeared in *Erkenntnis,* or maybe something that had appeared in *Erkenntnis* and something that Carnap hadn't published. At any rate, that particular version of it certainly appeared in English for the first time in *Readings in Philosophical Analysis.* And it seems to me that, although it's much shorter than Paul Horwich's book on truth, it is much more powerful. I say that without any detriment to Paul. He would be the last to claim to be a philosopher of Carnap's stature.

I'll use the term 'disquotationalism' for the Carnapian version of the theory and 'deflationism' for the Fregean version. On the Carnapian version, what's true and false are sentences, and sentences are marks and noises. (I take this language from Richard Rorty, but I think that Carnap would have no objection.) So we are supposed to say of a certain string of marks or noises that it's true. And we're told that to say of a string of marks or noises that it's true is just to assert that string of marks and noises. Now that version, of course, raises the question, what is it to assert marks and noises?

The only thing that remains is to continue doing what I've been doing at Harvard for thirty-five years, and at one place or another ever since I taught my first course in 1952 in Evanston: be a pedagogue. So I will do some teaching.

When Horwich (whom I use as a stalking horse in the third lecture) wrote his book on truth, he subscribed to an account of what it is to assert marks and noises—an account that was, in fact, exactly the account that Carnap would have given, though Horwich gave it elsewhere and not in that article. On Horwich's view at that time, to understand marks and noises is to be able to assign them a 'degree of probability' or perhaps a 'degree of assertibility '. Some of you, if you don't have tin ears, may already be beginning to wonder: how can marks and noises—say, the sequence of marks, 'There is a blackboard eraser on this table', regarded as a discontinuous range of patterns of ink on a page—be probable or assertible any more than being true? But Horwich explains, not in his book but in articles he published at the same time, that probability is something like a license to bet at certain odds. So we are supposed to have dispositions to assent to sentences—that is, presumably to mouth them—and, moreover, we have certain dispositions to bet at certain odds that we won't have to say, "I take it back," or something like that.

But this is precisely what Carnap would call assigning a degree of confirmation to a sentence. And indeed, in a review of Kripke's *Wittgenstein on Rules and*

Private Language, Horwich attributed this view to Wittgenstein. Now Horwich today will insist that this isn't his view anymore, and he once mildly criticized me for criticizing a view of his that he no longer holds—although he has published neither a retraction nor a sketch of what the replacement is going to be. In any case, Carnap's picture was quite clear.

Now, I would say that I'm not in the business, as a philosopher, of prohibiting you from talking in certain ways. I don't read Wittgenstein, either, as doing that. If you want to say of a sentence, in certain circumstances, that it's true, then OK: go ahead—provided that you recognize, at least, that sentences are only true or false under particular understandings. But presumably neither Horwich nor Carnap would object to that. Although Carnap might say, "I'm idealizing by assuming a language in which every sentence has one and only one fixed understanding."

But the model of an understanding of a sentence is functionalist: it's a disposition of a speaker, conceived of as if the speaker were a computer, to behave in certain ways or to lay certain wagers in response to certain stimulations. It's a methodologically solipsist picture. And ultimately, any methodologically solipsist picture will fail to do justice to the fact—which seems to be doubted only by French philosophers and people in English and French departments—that there is such a thing as representing the world and not just producing bets in response to inputs at the surface of your body.

In Frege's version, what are true and false primarily are judgments. And he denies that truth is a property—some universal that is wholly present in each true sentence or each true judgment. One way of making the judgment that there is a blackboard eraser on this table is to write this English sentence, another being to utter the corresponding noise; if I could recall the German phrase for "blackboard eraser," I could make that same judgment in German. And sometimes I think in German when I've been in Germany for a while. Or I think in Hebrew rather than in English. So I could make that judgment without either using or thinking the English sentence. Some people use, instead of 'judgment', the word 'statement', others 'proposition'; Wittgenstein in the *Philosophical Investigations* uses the word '*Satz*' in deliberate defiance, I think, of the doctrine that it must be either a sentence or something abstract and wholly distinct from the sentence. If I say of the judgment that there is a blackboard eraser on this table that it is true, I am not saying of one object, 'the judgment', that it has a property. On that view, then, whenever I think of a judgment that it is true I am making a meta-judgment about the original judgment. Whereas Frege wants to say that the relation between truth and judging is more intimate than that. (There's a good paper on this by Thomas Ricketts, by the way.) It is, rather, that when I say that it's true that there is a blackboard eraser on this table I'm judging that there is a blackboard eraser on this table. The subject is not the judgment; the subject is the blackboard eraser just as much as if I had only said, "There's a blackboard eraser on this table." That, at a certain superficial level, agrees with the other theory; that is, deflation and disquotation have a similar story to this extent.

Now the difference, I would say, is this. In the Fregean picture, judgments are not conceived of as corresponding to the world—it's rather a big thing to correspond to, especially if by 'the world' you mean 'the universe'—or even some piece of the world, or some peculiar entity in the world called, 'the fact that there's a

blackboard eraser on this table'. Rather, the judgment is intrinsically about the blackboard eraser, and the table, and the geometrical relation of 'being on.'

But what does that "intrinsic" talk come to? That sounds like mystery talk. Really, all it comes to is this: to be able to judge, to do what we call "judging," that there is a blackboard eraser on this table, you must have certain world-involving abilities. I would also speak here of language-involving abilities. I would defer to Warren Goldfarb's knowledge of Frege's texts here, but I think that this is something that Frege neither affirms nor denies. I don't think that he discusses the issue. But I see nothing intrinsically incompatible, in the Fregean view, with the claim that, at least for most judgments that human beings make—certainly judgments like the judgment that there is a blackboard eraser on this table—the capacity to make such judgments at all presupposes the sorts of skills that a speaker of the language comes to possess as he gains mastery of that language. And those skills involve such things as blackboard erasers, tables, and geometrical relations, and not only hypothetical events in the brain conceived of as a computer. Although today we don't even know, as I heard David Hubel say last week, how memory is laid down in the brain, or how memory is laid down in individual cells. It's rather a mystery, for the proteins in those cells are recycled, various things happen to them, and so on: how we manage to have stable memories is something that we still don't know. And yet we are happily babbling away about whether the brain is a computer.

Thus, in the Fregean story, supplemented in this somewhat Wittgensteinian way, there is the idea that using words is a world-involving thing. Wittgenstein speaks at times of 'methods of projection', which is, in a sense, representation—which is just what postmodernism denies ever exists. Now I think that Wittgenstein in one place uses the example of the phrase "blue sofa," and he says that you could of course say that the words "blue sofa" correspond in a particular context to a particular blue sofa. But Wittgenstein would not say that if I say, "There are no books in this room," then that sentence 'corresponds to' something called 'the absence of books in the room'. But there is nothing in talk of methods of projection, or of world-involving abilities, or of the idea that words do sometimes correspond to things, to force one to think that if some words, in some situations, can be meaningfully said to correspond to some particular things, then there must be one correspondence, one and the same correspondence, in every case where we can think of words as corresponding to particular things. Let alone that all words that can be meaningfully used, including whole sentences, correspond to particular things, even if we have to invent abstract things like 'the nonexistence of books in

In a discussion in which someone says, "I believe in correspondence truth; so-and-so rejects correspondence truth," there is always this unquestioned assumption that either it's all correspondence or no correspondence.

this room'.

That is, in a discussion in which someone says, "I believe in correspondence truth; so-and-so rejects correspondence truth," there is always this unquestioned assumption that either it's all correspondence or no correspondence; and, moreover, that it is one and the same correspondence always, or no correspondence.

So now I've completed my little task of saying something about the third Dewey Lecture.

Now another little bit of pedagogy. I have been talking a good deal about the unhelpfulness, to put it mildly—the nonsensicality—of sense datum talk, particularly when 'sense data' are said to be identical with neural processes. I have been looking at John Searle's *Minds, Brains, and Science,* an old book. He says in the first chapter that "the smell of the rose is a rate of neural firings." There you have the whole 'Cartesian-cum-materialist' picture in one sentence: the smell of the rose is a rate of neural firings. Now one thing that I have stressed is that (even if you aren't bothered by 'appearance' talk) there are no scientific objects that have the same identity conditions as appearances.

Consider that statement of Searle's: the smell of the rose is a rate of neural firings. First of all, it's ambiguous. Let's assume that the notion of token-identity of events is all right (I say in the second Dewey Lecture that all the existing definitions of token-identity for events have such utterly counterintuitive consequences that we're just better off not talking that way). But for the moment allowing talk of token-identity of events: is Searle saying that the whiteness of this paper, or the sense-datum or quale of the whiteness of this paper, is a rate of neural firings? Does he mean that this token event of its appearing white to me at this very instant is identical with the token event of particular neurons firing at a certain rate at that time? Let's grant that that might be true, if we knew what we meant by 'token-identity', and apart from other objections. Or does he mean that the property—the way something can be, or in this case the way a person can be—of its appearing to a person in that way, of having that experience of white, is a property of the form "having such-and-such neurons firing in such-and-such location at such-and-such rate"?

We are committed to an open plurality of ways of describing, ways of conceiving, ways of talking, ways of thinking: that, if you like, is pluralism.

Now, Searle has to mean the latter. For he compares the situation with the discovery that liquidity is explainable in terms of the properties of the water molecules, and that solidity is explainable in terms of the properties of a crystal; he even says with respect to intentionality that it is a consequence of his view that just as we can define—empirically and not analytically—liquidity or solidity, we may someday be able to define intentionality or the smell of the rose or the white of the chalk in terms of the properties of neuron firings. So his is clearly the view that one psycho-

logical attribute will be discovered to be a certain physical attribute.

And of course, I pointed out that no matter where you draw the boundaries for the rate of neuron firings—set them where you will, or legislate or just posit them if you like (as Nelson Goodman proposes to do, though not in a physicalist context)—the consequence will be that if the rate is just a little bit higher than the highest rate you allow, or just a little bit lower than that highest rate (even by one per second), you will have to say that you don't have that appearance: you have another one. So there will be two appearances such that no one could possibly tell them apart, even by the transitivity test. There won't be a third appearance that you could tell from one and not from the other, not if it is a difference of 1/10 or 1/100 second. Physical entities do not have the fuzziness that appearances have.

At this point, both of these examples—that from the theory of truth and that from the philosophy of psychology—begin to connect with two of the larger themes in my work in the last roughly twenty years. On the one hand, there is my pluralism. We have, to use the Wittgensteinian jargon, lots of language-games. (Although I think of language-games for Wittgenstein not as parts of language. He announces right away that he will use the phrase in a number of different ways, but very often they are rather models for parts of language, such that, as Wittgenstein says, you are to think both about their similarities to natural language and about their dissimilarities.) Pluralism has been a theme in my work, not always made with reference to Wittgenstein. I find it in James and elsewhere. But it has been a consistent theme in my work for almost twenty-five years: the language-games that we call "scientific language," or at least the ones that we paradigmatically think of, are insufficient to describe all of reality. There is no one language-game, no one group of language-games, of which you can say "All of reality is describable in terms of these, these constitute"—to use Quine's language, which I have criticized—"our first-grade conceptual system, and everything else is our second-grade conceptual system." Quine says our "second-grade system" is absolutely essential, of course. You have to use it when you take the agent point of view. It's indispensable in life. But you only need the "first-grade conceptual system" when you want to "limn the true and ultimate structure of reality."

It is in general an error to try to reduce one of our language-games to anything that looks on the surface like a very different one. Generally, if they look different on the surface, then they really are different.

This talk of "limning the true and ultimate structure of reality" I believe to be nonsense. We are committed to an open plurality of ways of describing, ways of conceiving, ways of talking, ways of thinking: that, if you like, is pluralism. It's also connected with anti-scientism. But I don't like to put it that way, for that sounds like anti-science. And I have always respected science enormously. In fact, I respect it too much to confuse it with science fiction. And one of the themes of this course

is that when you become scientistic in philosophy, you inevitably confuse science with science fiction, to the detriment of both science and philosophy.

Take the experience case. If you say that we need appearance talk, we need brain talk, we need many other kinds of talk, then you must see that there is no reduction here. Then of course you might respond, "But then there is a mystery: the mystery of emergence. Where do appearances emerge? Where does consciousness emerge?" I have argued (and this is perhaps the most Wittgensteinian strain in my recent work, for example, in my Royce Lectures) that that question itself only makes sense because we think, or we fall into the enormously seductive error of thinking, that we can imagine a world in which it had not emerged. We think that there could be a world like this one where all the people were physically just the same, but they were "automatic sweethearts" (in James's wonderful metaphor, which also occurs, by the way, in a story of E.T.A. Hoffmann's), or "empty hulls," as David Albert might put it. Maybe in this world consciousness only emerged once, only in you, Dear Reader! Maybe the rest of us, including this lecturer, are mere "empty hulls." That has only the appearance of sense. It makes a good story, but it relies heavily on the wholly unsound principle that whatever you can in any sense imagine is a conceivable possibility. That is a wholly unsound principle to rely upon anywhere in discussions of possibility.

This is not to say that I have an a priori proof that that sort of talk could never be given a sense. You won't find such a proof in my work; and I don't think that you will find such a proof, or an attempt at such a proof, in Wittgenstein's work either. All I say is that it hasn't been given a sense up to now. *Punkt.* Again, I don't ask you necessarily to buy that, but I am putting that forward as something I have defended.

But pluralism is not the only issue here. The other issue, I think, is connection. It is in general an error to try to reduce one of our language-games to anything that looks on the surface like a very different one. Generally, if they look different on the surface, then they really are different. It is the rare case when that is only an appearance. And in a way, it is an error to think, "If these language-games—that of talk about appearances and sensations, and that of neurology, and that of behavior—are all different, then there are no connections." What I am suggesting is that analytic philosophy, starting with logical positivism and perhaps earlier, valorized one kind of connection too much. It valorized strict equivalence: biconditionals, definitions, finding out that p if and only if q. Such connections are rare. But "softer" connections—"When we conceptualize in this way, we rely on the availability of this other form of conceptualization"—are all over the place. And part of the impression, post-Wittgenstein, that all that is left is the end of philosophy, or quietism, or saying nothing, is the failure to see any interest in the enormous range of connections, connections among all our different language-games, which are still largely unexplored. For we are still recycling positions in philosophy that were familiar to Kant before he wrote the first *Critique*, and we are only interested in what might support one of those.

Of course, the form of pluralism that I have most strongly advocated involves the recognition of what I call "the collapse of the fact-value dichotomy." I know that some of you in this class have been surprised to hear me even question the view that "values are subjective, and that's the end of the story." But that is

indeed a view that I have questioned and have argued to be incoherent. One of the ways that I have argued—and I'm building here on the work of John Dewey—is by insisting on the ubiquity of value. Dewey, in a letter to James of 1903, says, in effect, "People think that value is something that occupies some little corner of experience, some little area here. But I think that it is absolutely ubiquitous." And I urge that we have cognitive values, for example, coherence. My examples often come from big scientific theories. Last week, at a meeting of the American Philosophical Society in Philadelphia at which I lectured about this, the scientists in the audience, who were mostly Nobel Prize winners, included a number of people who were enthusiastic about this claim; they said, "Of course we judge theories on the basis of things like beauty." Gerald Holton has produced a series of examples, all connected with the Special Theory of Relativity, starting with Planck's answer to the criticism, "Why did you accept Einstein's theory? We have Lorentz transformations, and we have Poincaré. Why Einstein?" And Planck replied, "Es ist mir einfach sympathischer." It's just more *simpatico*. And when I reminded Holton of this story, he produced two other scientists who were originally opponents of the Special Theory until after it was formulated by Minkowski in a way that really brought out its elegance, and were won over. One of them said that the theory is so beautiful that it has to be true.

If you take two events six minutes apart on Mars, and you tell me that there is no fact of the matter as to which of those is happening right now, that sounds nuts. And yet the beauty of Einstein's theory overrode the certainty of that a priori truth.

But I am not saying that it is good pragmatist, or any other, methodology to say that you should accept a theory on beauty alone. Although we often refuse to test a theory just because it is ugly. For you cannot test every possible theory. You cannot even test every possible unrefuted theory. There is the example of Whitehead's theory of gravitation, which was never tested until the seventies, long after the rival, Einstein's theory of gravitation, had been accepted—simply because it was so ugly!

But it is not only in the areas of big theories, so-called scientific revolutions, that values of coherence, elegance, and beauty can sometimes outweigh long-standing, millennia-old judgments of what is a priori the case. For millennia we believed that if you have an event on Earth and an event on Mars, then either they happen at the same time or they don't. Period. We believed that, apart from very fine discriminations, like which horse won a race, there is a fact of the matter as to which happened first. If you take two events six minutes apart on Mars, and you tell me that there is no fact of the matter as to which of those is happening right now, that sounds nuts. And yet the beauty of Einstein's theory overrode the certainty of that a priori truth. (And it is one of Kant's a priori truths.)

But there are also judgments of coherence on a much more mundane level. Such judgments are involved when I decide which of my memories to trust, for example. I talked at the same meeting last week with a scientist who said, "My real problem isn't with the mathematical formula; the real problem is deciding which data to trust." This is a question of coherence. Now here again, if you are fixed with the view, which I think dominates a lot of analytic metaphysics, that the only predicates we can take seriously have to do either with what causes what or with composition (or with both composition and causation), then what can you say about the role of a predicate like 'is an incoherent theory', 'is a Rube Goldberg job', 'is ad hoc'? These phrases can be descriptive: certainly, when I say of a theory that it has "artificial assumptions," I am making a description. If I use words from logic—if I say "This is a valid proof," or "This is an invalid proof; there's a fallacy at line five"—I am describing something. But I am not speaking the language of efficient causation and composition. And in many ways, these cognitive values—coherence, plausibility, ad hoc-ness and so on—work much more like 'valid' and 'invalid'. In fact, in the wide and loose sense of 'logic' that J.S. Mill and Dewey used, they are logical words; they belong to the theory of inquiry. And again, what Wittgenstein called "the philosopher's '*must*'" makes us unable to see this. The philosopher's '*must*' always functions as a pair of blinkers.

To return to the collapse of the fact-value dichotomy: I talked in my last lecture, in connection with the case of ethical value, about Dewey. His answer to the questions, "How can you responsibly discuss ethical values? How can you responsibly discuss a question like whether a society should conceive itself, in Rawlsian terms, as a cooperative venture among free and equal citizens?" Part of the answer, of course, is: look at the reasons that people give for denying that people should be treated as equals. What reasons have people given for saying that women should not be treated as equals, that blacks should not be treated as equals, and so on? It is not as if these things occurred in a vacuum. Perhaps the worst thing about subjectivism is that it is a reasoning-stopper. This is something that Michele Moody-Adams, who is a black woman philosopher, by the way, argues in *Fieldwork in Familiar Places*, brought out by Harvard University Press two years ago.

There are two other points about this whole question of subjectivism versus objectivism. First, subjectivist views have enormous impact in our culture. Not only, as I mentioned in *Reason, Truth, and History*, do they constantly appear at cocktail parties. (Even if they were confined to cocktail parties, it would still be serious; if it appeared at most all of them, it would mean something politically.) For example, think of how much influence on economic theory has been exerted by the idea that there is nothing rational about valuation except in the sense in which subjective preferences can be rational—how it is built into economic theory. I recently discovered that a very early article by Amartya Sen, who was a member of this department as well as a Nobel Prize winning economist, attacked the fact-value dichotomy; that is something he has been concerned with all his life, for this reason. Think of how the role, especially in the so-called Chicago School, of that kind of subjectivism is now beginning to spread through the law.

But if any philosophical issue deserves to be called a perennial issue in philosophy, that one surely does. For it is all over the place in Plato and Aristotle. But you might say, "Yes, but Plato was unconvincing. He gets out of it with the theory

positing these Forms that we can't believe in." First of all, it isn't clear that Plato stuck with those Forms either. Some of the earliest and most serious criticisms of the Theory of Forms also come from Plato. And in the *Nicomachean Ethics*, Aristotle deals with the same issue. And transcendentalism plays very little role in that work, at most in Book X. And even there I think that its role is generally exaggerated. And there is a long question about the consistency of Books I-IX with Book X.

Is there anything that I have left out of this lecture? Of course, one doesn't do philosophy only because one hopes that it will make some difference in the world. One also does it for the pure joy of it. And those are not incompatible. One of Plato's dialogues begins with a kind of short replay of the *Apology*. It is set in a city other than Athens, and it begins with a speaker who has come from Athens and is telling the people in this other Greek town what Socrates's death was like. And I actually read that part of it in Greek; my Greek is very slow and rusty, but I did it. And the visitor from Athens describes how they were talking philosophy with Socrates after he had drunk the hemlock. And the speaker uses a Greek phrase which is almost morpheme-for-morpheme synonymous with the English expression "to be torn:" "We were torn, because on the one hand it was such a pleasure. And on the other hand we knew that he was dying." And of course, philosophy is a great pleasure, and a pleasure that I hope to continue having for a long time. And as the example of Socrates shows, experiencing it as a pleasure and doing it as a pleasure are not incompatible with being aware of your responsibilities to society and your responsibilities to your own self-betterment.

The best and also the worst closing lecture of a career at Harvard was given by Wilfred Cantwell Smith, for whom there will be a memorial service, I think, tomorrow. It was the worst because Smith's delivery was so boring that my eyes keep falling shut. But, at the same time, I kept having the experience of realizing that the sentence I had just heard was one of the most beautiful sentences that I had ever heard in my life. I think that Smith deliberately would read out his lectures this way to avoid any hint of rhetoric. But he of course was a professor of religion; he chaired the Committee on the Study of Religion here at Harvard. And at the close of his lecture he said, "I'm not saying that religion is a good thing. I'm saying that it's a great thing. It can make you better or it can make you much worse. But it means that you take the question of how to live seriously." And if I were to mimic that, I would say that philosophy isn't a good thing. It's a great thing. It can lead to wonderful things, and it can lead to terrible things. But it means—to modify Smith's sentence—that you take the responsibility of trying to think deeply and with integrity seriously. φ