

Paul M. Churchland  
THE ANTI-REALIST EPISTEMOLOGY  
OF VAN FRAASSEN'S  
THE SCIENTIFIC IMAGE\*

At several points in the reading of van Fraassen's book, I feared I would no longer be a realist by the time I completed it. Fortunately, sheer doxastic inertia has allowed my convictions to survive its searching critique, at least temporarily, and as of today, van Fraassen and I still hold different views. I am a scientific realist, of unorthodox persuasion, and van Fraassen is a constructive empiricist, whose persuasions currently define the doctrine. I assert that global excellence of theory is the ultimate measure of truth and ontology at all levels of cognition, even at the observational level. Van Fraassen asserts that descriptive excellence at the observational level is the only genuine measure of any theory's truth, and that one's acceptance of a theory should create no ontological commitments whatever beyond the observational level.

Against his first claim I will maintain that observational excellence or 'empirical adequacy' is only one epistemic virtue among others, of equal or comparable importance. And against his second claim I will maintain that the ontological commitments of any

\*This paper is an expanded version of a presentation delivered at the annual meetings of the Canadian Philosophical Association (in Halifax, May, 1981), in a session devoted to a critical discussion of van Fraassen [5], and Churchland [1]. Van Fraassen's contribution to that session has appeared, as a critical notice of [1], in [6].

I should like to thank Hartry Field, Michael Stack, Bas van Fraassen, Clark Glymour, Barney Keaney, Stephen Stich, and Patricia Churchland for helpful discussion of the issues here addressed.

Reprinted by permission from *Pacific Philosophical Quarterly*, Vol. 63, no. 2, July 1982.

theory are wholly blind to the idiosyncratic distinction between what is and what is not humanly observable, and so should be our own ontological commitments. Criticism will be directed primarily at van Fraassen's *selective* scepticism in favor of observable ontologies over unobservable ontologies; and against his view that the superempirical theoretical virtues (simplicity, coherence, explanatory power) are merely pragmatic virtues, irrelevant to the estimate of a theory's truth. My aims are not merely critical, however. Scientific realism does need reworking, and there are good reasons for moving it in the direction of van Fraassen's constructive empiricism, as will be discussed in the closing section of this paper. But those reasons do not support the sceptical theses at issue.

\* \* \*

Before pursuing our differences, it will prove useful to emphasize certain convictions we share. Van Fraassen is already a scientific realist in the minimal sense that he interprets theories literally and he concedes them a truth-value. Further, we agree that the observable/unobservable distinction is entirely distinct from the non-theoretical/theoretical distinction, and we agree as well that all observation sentences are irredeemably laden with theory.

Additionally, I absolutely reject many sanguine assumptions common among realists. I do not believe that on the whole our beliefs must be at least roughly true; I do not believe that the terms of 'mature' sci-

ences must typically refer to real things; and I very much doubt that the Reason of *homo sapiens*, even at its best and even if allowed infinite time, would eventually encompass all and/or only true statements.

This scepticism is born partly from a historical induction: so many past theories, rightly judged excellent at the time, have since proved to be false. And their current successors, though even better founded, seem but the next step in a probably endless and not obviously convergent journey. (For a most thorough and insightful critique of typical realist theses, see the recent paper by Laudan [4].)

Evolutionary considerations also counsel a healthy scepticism. Human reason is a hierarchy of heuristics for seeking, recognizing, storing, and exploiting information. But those heuristics were invented at random, and they were selected for within a very narrow evolutionary environment, cosmologically speaking. It would be miraculous if human reason were completely free of false strategies and fundamental cognitive limitations, and doubly miraculous if the theories we accept failed to reflect those defects.

Thus some very realistic reasons for scepticism with respect to any theory. Why then am I still a scientific realist? Because these reasons fail to discriminate between the integrity of observables and the integrity of unobservables. If anything is compromised by these considerations, it is the integrity of theories generally. That is, of *cognition* generally. Since our observational concepts are just as theory-laden as any others, and since the integrity of those concepts is just as contingent on the integrity of the theories that embed them, our observational ontology is rendered *exactly as dubious* as our non-observational ontology.

This parity should not seem surprising. Our history reveals mistaken ontological commitments in both domains. For example, we have had occasion to banish phlogiston, caloric, and the luminiferous aether from our ontology, but we have also had occasion to banish witches, and the starry

sphere that turns about us daily. And these latter items were as 'observable' as you please.

Since these sceptical considerations are indifferent to the distinction between what is and is not observable, they provide no reason for resisting a commitment to unobservable ontologies *while allowing* a commitment to observable ontologies. The latter appear as no better off than the former. For me then, the empirical success of a theory to be true, and for accepting its overall ontology. The inference from success to truth should no doubt be tempered by the sceptical considerations adduced, but the inference to *unobservable* ontologies is not rendered *selectively* dubious. Thus I remain a scientific realist. My realism is highly circumspect, but the circumspection is uniform for unobservables and observables alike.

Perhaps I am wrong in this. Perhaps we should be selectively sceptical in the fashion van Fraassen recommends. Does he have other arguments for refusing factual belief and ontological commitment beyond the observational domain? Indeed, he does. In fact, he does not appeal to historical induction or evolutionary humility at all. These are *my* reasons for scepticism (and they will remain, even if I manage to undermine van Fraassen's). They have been introduced here to show that, while there are some powerful reasons for scepticism, those reasons do not place unobservables at a selective disadvantage.

Very well, what are van Fraassen's reasons for scepticism? They are very interesting. To summarize quickly, he does a compelling job of deflating certain standard realist arguments (from Smart, Sellars, Salmon, Boyd, and others) to the effect that, given the aims of science, we have no alternative but to bring unobservables (not just into our calculations, but) into our literal ontology. He also argues rather compellingly that the superempirical virtues, such as simplicity and comprehensive explanatory power, are at bottom merely

pragmatic virtues, having nothing essential to do with any theory's truth. This leaves only empirical adequacy as a genuine measure of any theory's truth. Roughly, a theory is empirically adequate if and only if everything it says about *observable* things is true. Empirical adequacy is thus a necessary condition on a theory's truth.

However, claims van Fraassen, the truth of any theory whose ontology includes unobservables is always radically underdetermined by its empirical adequacy since a great many logically incompatible theories can all be empirically equivalent. Accordingly, the inference from empirical adequacy to truth now appears presumptuous in the extreme, especially since it has just been disconnected from additional selective criteria such as simplicity and explanatory power, criteria which might have reduced the arbitrariness of the particular inference drawn. Fortunately, says van Fraassen, we do not need to make such wanton inferences since we can perfectly well understand science as an enterprise that never really draws them. Here we arrive at his positive conception of science as an enterprise whose sole intellectual aims are empirical adequacy and the satisfaction of certain human intellectual needs.

The central element in this argument is the claim that, in the case of a theory whose ontology includes unobservables, its empirical adequacy underdetermines its truth. (We should notice that in the case of a theory whose ontology is completely free of unobservables, its empirical adequacy does not underdetermine its truth: in that case, truth and empirical adequacy are obviously identical. Thus van Fraassen's *selective* scepticism with respect to unobservables.) That is, for any theory T inflated with unobservables, there will always be many other such theories incompatible with T, but empirically equivalent to it.

In my view, the notions of "empirical adequacy" and its cognate relative term "empirically equivalent" are extremely thorny notions of doubtful integrity. If we attempt to explicate a theory's "empirical

content" in terms of the observation sentences it entails (or entails-if-conjoined-with available background information, or with possible future background information, or with possible future theories), we generate a variety of notions which are variously empty, context-relative, ill-defined, or incompatible with the claim of underdetermination. Van Fraassen is entirely aware of these difficulties and proposes to avoid them by giving the notions at issue a model-theoretic rather than a syntactic explication. I am unconvinced that this improves matters decisively (on this issue see Wilson [7]). But let me sidestep the issue for now, since the matter is difficult and there is a simpler objection to be voiced.

The empirical adequacy of any theory is itself something that is radically underdetermined by any evidence conceivably available to us. Recall that, for a theory to be empirically adequate, what it says about observable things must be true—all observable things, in the past, in the indefinite future, and in the most distant corners of the cosmos. But since any actual data possessed by us must be finite in its scope, it is plain that we here suffer an underdetermination problem no less serious than that claimed above. This is Hume's problem, and the lesson is that even observation-level theories suffer radical underdetermination by the evidence. Accordingly, theories about observables and theories about unobservables appear on a par again, so far as scepticism is concerned.

Van Fraassen thinks there is an important difference between the two cases, and one's first impulse is to agree with him. We are all willing to concede the existence of Hume's problem—the problem of justifying the inference to unobserved entities. But the inference to entities that are downright unobservable appears as a different and additional problem.

I do not see that it is. Consider the different reasons why entities or processes may go unobserved by us. First, they may go unobserved because, relative to our natural sensory apparatus, they fail to enjoy an

appropriate spatial or temporal *position*. They may exist in the Upper Jurassic Period, for example, or they may reside in the Andromeda Galaxy. Second, they may go unobserved because, relative to our natural sensory apparatus, they fail to enjoy the appropriate spatial or temporal *dimensions*. They may be too small, or too brief, or too large, or too protracted. Third, they may fail to enjoy the appropriate *energy*, being too feeble, or too powerful, to permit useful discrimination. Fourth and fifth, they may fail to have an appropriate *wavelength*, or an appropriate *mass*. Sixth, they may fail to 'feel' the relevant fundamental forces our sensory apparatus exploits, as with our inability to observe the background neutrino flux, despite the fact that its energy density exceeds that of light itself.

This list could be lengthened, but it is long enough to suggest that being spatially or temporally distant from our sensory apparatus is only one undistinguished way, one among many ways, in which an entity or process can fall outside the compass of human observation.

There is perhaps some point to calling a thing "observable" if it fails only the first test (spatio-temporal proximity), and "unobservable" if it fails any of the others. But that is only because of the contingent practical fact that one generally has somewhat more *control* over the spatio-temporal perspective of one's sensory systems than one has over their size, or reaction time, or mass, or wavelength sensitivity, or chemical constitution. Had we been less mobile—rooted to the earth like Douglas Firs, say—yet been more voluntarily plastic in our sensory constitution, the distinction between the 'merely unobserved' and the 'downright unobservable' would have been very differently drawn. It may help to imagine here a suitably rooted arboreal philosopher named (what else?) Douglas van Fraassen, who, in his sedentary wisdom, urges an anti-realist scepticism concerning the spatially very *distant* entities postulated by his fellow trees.

Admittedly, for any distant entity one can in principle always change the relative spa-

tial position of one's sensory apparatus so that the entity is observed: one can go to it. But equally, for any microscopic entity one can in principle always change the relative spatial *size* or *configuration* of one's sensory apparatus so that the entity is observed. Physical law imposes certain limitations on such plasticity, but so also does physical law limit how far one can travel in a lifetime.

The point of all this is that there is no special or additional problem about inferences to the existence of entities commonly called "unobservables." Such entities are merely those that go unobserved by us for reasons *other* than their spatial or temporal distance from us. But whether the 'gap' to be bridged is spatio-temporal, or one of the many other gaps, the logical/epistemological problem is the same in all cases: ampliative inference and underdetermined hypotheses. I therefore fail to see how van Fraassen can justify tolerating an ampliative inference when it bridges a gap of spatial distance, while refusing to tolerate an ampliative inference when it bridges a gap of, for example, spatial size. Hume's problem and van Fraassen's problem (or Duhem's problem) collapse into one.

Van Fraassen attempts to meet such worries about the inescapable ubiquity of speculative activity by observing that ". . . it is not an epistemological principle that one may as well hang for a sheep as for a lamb" ([5], p. 72). Agreed. But it is a principle of *logic* that one may as well hang for a sheep as for a sheep, and van Fraassen's lamb (empirical adequacy) is just another sheep.

Let me summarize. As van Fraassen sets it up, and as the instrumentalists set it up before him, the realist looks more gullible than the non-realist, since the realist is willing to extend belief beyond the observable, while the non-realist insists on confining belief within that domain. I suggest, however, that it is really the non-realists who are being the more gullible in this matter, since they suppose that the epistemic situation of our beliefs about observables is in some way superior to that of our beliefs about unobservables. But in fact their epis-

exactly as dubious as their non-observational cousins. Their causal history is different perhaps, but not their credibility.

Simply to hold *fewer* beliefs from a given set is of course to be less adventurous, but it is not necessarily to be applauded. One might decide to relinquish all one's beliefs save those about objects weighing less than 500 kg, and perhaps one would then be logically safer. But in the absence of some relevant epistemic difference between one's beliefs about such objects and one's beliefs about other objects, this is perversity, not parsimony.

\* \* \*

Let me now try to address the question of whether the theoretical virtues such as simplicity, coherence, and explanatory power are *epistemic* virtues genuinely relevant to the estimate of a theory's truth, as tradition says, or merely *pragmatic* virtues, as van Fraassen urges. His view promotes empirical adequacy, or evidence of empirical adequacy, as the only genuine measure of a theory's truth, the other virtues (insofar as they are distinct from these) being cast as purely pragmatic virtues, to be valued only for the human needs they satisfy. Despite certain compelling features of the account of explanation that van Fraassen provides, I remain inclined towards the traditional view.

My reason is simplicity itself. Since there is no way of conceiving or representing "the empirical facts" that is completely independent of speculative assumptions, and since we will occasionally confront theoretical alternatives on a scale so comprehensive that we must also choose between competing modes of conceiving what the empirical facts before us *are*, then the epistemic choice between these global alternatives cannot be made by comparing the extent to which they are adequate to some common touchstone, 'the empirical facts.' In such a case, the choice must be made on the comparative global virtues of the two global alternatives,  $T_1$ -plus-the-observational-evidence-therein-

al-evidence-therein-(differently)-construed. That is, it must be made on superempirical grounds such as relative coherence, simplicity, and explanatory unity. In such cases, "empirical adequacy" becomes just one dimension of coherence.

Such cases as these are reminiscent of Carnap's 'external' questions, and it may be that van Fraassen, like Carnap, does not regard them as factual questions, but as essentially pragmatic questions. I would disagree, since I regard so-called 'external' questions as arrayed on a smooth continuum with 'internal' (i.e., factual) questions. The arguments are presented elsewhere ([1], sections 7 and 10), however, so I shall not repeat them here.

As I see it then, values such as ontological simplicity, coherence, and explanatory power are some of the brain's criteria for recognizing information, for distinguishing information from noise. And I think they are even more fundamental values than is 'empirical adequacy,' since collectively they can overthrow an entire conceptual framework for representing the empirical facts. Indeed, they even dictate how such a framework is constructed by the questing infant in the first place. One's observational taxonomy is not 'read off' the world directly; rather, one comes to it piecemeal, and by stages, and one settles on that taxonomy which finds the greatest coherence and simplicity in the world, and most and the simplest lawful connections.

I can bring together my protective concerns for unobservables and for the superempirical virtues by way of the following thought experiment. Consider a man for whom absolutely *nothing* is observable. All of his sensory modalities have been surgically destroyed, and he has no visual, tactile, or other sensory experience of any kind. Fortunately, he has mounted on top of his skull a microcomputer fitted out with a variety of environmentally-sensitive transducers. The computer is connected to his association cortex (or perhaps the frontal lobe, or Wernicke's area) in such a way as to cause in him

a continuous string of singular beliefs about his local environment. These "intellectual intuitions" are not infallible, but let us suppose that they provide him with much the same information that our perceptual judgments provide us.

For such a person, or for a society of such persons, the *observable* world is an empty set. There is no question, therefore, of their evaluating any theory by reference to its 'empirical adequacy,' as characterized by van Fraassen (i.e., isomorphism between some observable features of the world and some 'empirical substructure' of one of the theory's models). But such a society is still capable of science, I assert. They can invent theories, construct explanations of the facts-as-represented-in-past-spontaneous-beliefs, hazard predictions of the facts-as-represented-in-future-spontaneous-beliefs, and so forth. In principle, there is no reason they could not learn as much as we have. (cf. Feyerabend [3])

But it is plain in this case that the global virtues of simplicity, coherence, and explanatory unification are what *must* guide the continuing evolution of their collected beliefs. And it is plain as well that their ontology, whatever it is, must consist entirely of *unobservable* entities. To invite a van Fraassenean disbelief in unobservable entities is in this case to invite the suspension of all beliefs beyond tautologies! Surely reason does not require them to be so abstemious.

It is time to consider the objection that those aspects of the world which are successfully monitored by the transducing microcomputer should count as 'observables' for the folk described, despite the lack of any appropriate field of internal sensory qualia to mediate the external circumstance and the internal judgment it causes. Their tables-and-chairs ontology, as expressed in their spontaneous judgments, could then be conceded legitimacy.

I will be the first to accept such an objection. But if we do accept it, then I do not see how we can justify van Fraassen's selective scepticism with respect to the wealth of

'unobservable' entities and properties reliably monitored by *our* transducing measuring instruments (electron microscopes, cloud chambers, chromatographs, etc.). The spontaneous singular judgments of the working scientist, at home in his theoretical vocabulary and deeply familiar with the measuring instruments to which his conceptual system is responding, are not worse off, causally or epistemologically, than the spontaneous singular judgments of our transducer-laden friends. If scepticism is to be put aside above, it must be put aside here as well.

My concluding thought experiment is a complement to the one just outlined. Consider some folk who observe, not less of the world than we do, but more of it. Suppose them able to observe a domain normally closed to us: the micro-world of virus particles, DNA strands, and large protein molecules. Specifically, suppose a race of humanoid creatures each of whom is born with an electron microscope permanently in place over his left 'eye.' The scope is biologically constituted, let us suppose, and it projects its image onto a human-style retina, with the rest of their neurophysiology paralleling our own.

Science tells us, and I take it that van Fraassen would agree, that virus particles, DNA strands, and most other objects of comparable dimensions count as observable entities for the humanoids described. The humanoids, at least, would be justified in so regarding them and in including them in their ontology.

But we humans may not include such entities in our ontology, according to van Fraassen's position, since they are not observable with our unaided perceptual apparatus. We may not include such entities in our ontology *even though we can construct and even if we do construct electron microscopes of identical function, place them over our left eyes, and enjoy exactly the same microexperience as the humanoids.*

The difficulty for van Fraassen's position, if I understand it correctly, is that his position requires that a humanoid and a scope-

epistemic attitudes towards the microworld, even though their causal connections to the world and their continuing experience of it be identical: the humanoid is required to be a realist with respect to the microworld, and the human is required to be an anti-realist (i.e., an agnostic) with respect to the microworld. But this distinction between what we and they may properly embrace as real seems to me to be highly arbitrary and radically under-motivated. For the only difference between the humanoid and a scope-equipped human lies in the *causal origins* of the transducing instruments feeding information into their respective brains. The humanoid's scope owes its existence to information coded in his genetic material. The human's scope owes its existence to information coded in his cortical material, or in technical libraries. I do not see why this should make any difference in their respective ontological commitments, whatever they are, and I must decline to embrace any philosophy of science which says that it must.

\* \* \*

I now turn from critic of van Fraassen's position to advocate. One of the most central elements in his view seems to me to be well-motivated and urgently deserving of further development. As he explains in his introductory chapter, his aim is to reconceive the relation of theory to world, and the units of scientific cognition, and the virtue of those units when successful. He says,

I use the adjective 'constructive' to indicate my view that scientific activity is one of construction rather than discovery: construction of models that must be adequate to the phenomena, and not discovery of truth concerning the unobservable. ([5], p. 5)

The traditional view of human knowledge is that the unit of cognition is the sentence or proposition, and the cognitive virtue of such units is truth. Van Fraassen rejects this overtly linguistic guise for his

theory as a set of models (rather than as a set of sentences), and he sees empirical adequacy (rather than truth) as the principal virtue of such units.

Though I reject his particular reconception, and the selective scepticism he draws from it, I think the move away from the traditional conception is entirely correct. The criticism to which I am inclined is that van Fraassen has not moved quite far enough. Specifically, if we are to reconsider truth as the aim or product of cognitive activity, I think we must reconsider its applicability right across the board, and not just in some arbitrarily or idiosyncratically segregated domain of 'unobservables.' That is, if we are to move away from the more naive formulations of scientific realism, we should move in the direction of *pragmatism* rather than in the direction of a positivistic instrumentalism. Let me elaborate.

When we consider the great variety of cognitively active creatures on this planet—sea slugs and octopi, bats, dolphins, and humans; and when we consider the ceaseless reconfiguration in which their brains or central ganglia engage—adjustments in the response potentials of single neurons made in the microsecond range, changes in the response characteristics of large systems of neurons made in the seconds-to-hours range, dendritic growth and new synaptic connections and the selective atrophy of old connections effected in the day-upwards range; then van Fraassen's term 'construction' begins to seem highly appropriate. There is endless construction and reconstruction, both functional and structural. Further, it is far from obvious that truth is either the primary aim or the principal product of this activity. Rather, its function would appear to be the ever more finely tuned administration of the organism's *behavior*. Natural selection does not care whether a brain has or tends toward true beliefs, so long as the organism reliably exhibits reproductively advantageous behavior. Plainly there is going to be *some* connection between the faithfulness of the

the organism's behavior. But just as plainly, the connection is not going to be direct.

While we are considering cognitive activity in biological terms and in all branches of the phylogenetic tree, we should note that it is far from obvious that sentences or propositions or anything remotely like them constitute the basic elements of cognition in creatures generally. Indeed, as I have argued at length elsewhere ([1], chapter 5; [2]), it is highly unlikely that the sentential kinematics embraced by folk psychology and orthodox epistemology represents or captures the basic parameters of cognition and learning even in humans. That framework is part of a common-sense theory that threatens to be either superficial or false. If we are ever to understand the *dynamics* of cognitive activity, therefore, we may have to reconceive our basic unit of cognition as something other than the sentence or proposition, and reconceive its virtue as something other than truth.

Success of this sort on the descriptive/explanatory front would likely have normative consequences. Truth, as currently conceived, might cease to be an aim of science. Not because we had lowered our sights and reduced our epistemic standards, as van Fraassen's constructive empiricism would suggest, but because we had raised our sights, in pursuit of some epistemic goal even *more* worthy than truth. I cannot now elucidate such goals, but we should be sensible of their possible existence. The notion of "truth," after all, is but the central element in a normative *theory*, and *praxis* makes progress no less than *theoria*.

The notion of truth is suspect on purely metaphysical grounds anyway. It suggests straightaway the notion of the Complete and Final True Theory: at a minimum, the infinite set of all true sentences. Such a theory would be, by epistemic criteria, the best theory possible. But nothing whatever guarantees the existence of such a unique theory. Just as there is no largest positive integer, it may be that there is no best the-

ory. There may be that for any theory whatsoever, there is always an even better theory, and so ad infinitum. If we were thus unable to speak of the set of all true sentences, what sense could we make of truth sentence-by-sentence?

These considerations do invite a 'constructive' conception of cognitive activity, one in which the notion of truth plays at best a highly derivative role. The formulation of such a conception, adequate to all of our epistemic criteria, is the outstanding task of epistemology. I do not think we will find that conception in a model-theoretic version of positivistic instrumentalism, nor do I think we will find it quickly. But the empirical brain begs unravelling, and we have plenty of time.

Finally, there is a question put to me by Stephen Stich. If ultimately my view is even more sceptical than van Fraassen's concerning the relevance or applicability of the notion of truth, why call it scientific *realism* at all? For at least two reasons. The term "realism" still marks the principal contrast with its traditional adversary, positivistic instrumentalism. Whatever the integrity of the notion of truth, theories about unobservables have just *as much* a claim to truth, epistemologically and metaphysically, as theories about observables. Second, I remain committed to the idea that there exists a world, independent of our cognition, with which we interact, and of which we construct representations: for varying purposes, with varying penetration, and with varying success. Lastly, our best and most penetrating grasp of the real is still held to reside in the representations provided by our best theories. Global excellence of theory remains the fundamental measure of rational ontology. And that has always been the central claim of scientific realism.

## REFERENCES

1. Churchland, Paul M. (1979), *Scientific Realism and the Plasticity of Mind* (Cambridge University Press).



2. — (1981), "Eliminative Materialism and the Propositional Attitudes," *The Journal of Philosophy*, vol. 78, no. 2, February.
3. Feyerabend, Paul K. (1969), "Science Without Experience," *The Journal of Philosophy*, vol. 66, no. 22.
4. Laudan, Larry (1981), "A Confutation of Convergent Realism," *Philosophy of Science*, vol. 48, no. 1, March.

5. van Fraassen, Bas C. (1980), *The Scientific Image* (Oxford University Press).
6. — (1981), "Critical Notice of Paul Churchland: Scientific Realism and the Plasticity of Mind," *Canadian Journal of Philosophy*, vol. 11, no. 3, September.
7. Wilson, Mark (1980), "The Observational Uniqueness of Some Theories," *The Journal of Philosophy*, vol. 77, no. 4, April.

## Hilary Putnam EXPLANATION AND REFERENCE\*

### I. GENERAL SIGNIFICANCE OF THE TOPIC

In this paper I try to contrast realist theories of meaning with what may be called 'idealist' theories of meaning. But a word of explanation is clearly in order.

There is no Marxist 'theory of meaning' but there are a series of remarks on the correspondence between concepts and things, on concepts, and on the impossibility of *a priori* knowledge in the writings of Engels (cf. Engels, 1959) which clearly bear on problems of meaning and reference. In particular, there is a passage† in which Engels makes the point that a concept may contain elements which are not correct. A contemporary scientific characterization of fish would include, Engels says, such properties as life under water and breathing through gills; yet lungfish and other anomalous spe-

cies which lack these properties are classified as fish for scientific purposes. And Engels argues, I think correctly, that to stick to the letter of the 'definition' in applying the concept *fish* would be bad science. In short, Engels contends that:

- (1) Our scientific conception (I would say 'stereotype') of a fish includes the property 'breathing through gills', but
- (2) 'All fish breath through gills' is not true! (and, *a fortiori*, not analytic).

I do not wish to ascribe to Engels an anachronistic sophistication about contemporary logical issues, but without doing this it is fair to say on the basis of this argument that Engels *rejects* the model according to which such a concept as *fish* provides anything like analytically necessary and sufficient conditions for membership in a natural kind. Two further points are of importance: (1) The fact that the concept 'natural kind *all* of whose members live under water, breath through gills, etc.' does not strictly fit the natural kind Fish does not mean that the concept does not *correspond* to the natural kind Fish. As Engels puts it, the

concept is not exactly correct (as a description of the corresponding natural kind) but that does not make it a *fiction*. (2) The concept is continually changing as a result of the impact of scientific discoveries, but that does not mean that it ceases to correspond to the same natural kind (which is itself, of course, also changing). Again, without attributing to Engels a sophisticated theory of meaning and reference, it is fair, I think, to restate the essential gist of these two points in the following way: concepts which are not strictly true of anything may yet refer to something; and concepts in different theories may refer to the same thing. Of these two points, the second is obvious for most realists; with a few possible exceptions (e.g., Paul Feyerabend), realists have held that there are successive scientific theories about the *same* things: about heat, about electricity, about electrons, and so forth; and this involves treating such terms as 'electricity' as *trans-theoretical* terms, as Dudley Shapere has called them (cf. Shapere, 1969), i.e., as terms that have the same reference in different theories. The first point is more controversial; the idea that concepts provide necessary and sufficient conditions for class membership has often been attacked but, nonetheless, constantly reappears. Without it, however, the other point is moot. Bohr assumed in 1911 that there are (at every time) numbers  $p$  and  $q$  such that the (one dimensional) position of a particle is  $q$  and the (one dimensional) momentum is  $p$ ; if this was part of the meaning of 'particle' for Bohr, and in addition, 'part of the meaning' means 'necessary condition for membership in the extension of the term', then electrons are *not* particles in Bohr's sense, and, indeed, there are *no* particles 'in Bohr's sense'. (And no 'electrons' in Bohr's sense of 'electron', etc.) None of the terms in Bohr's 1911 theory referred! It follows on this account that we cannot say that present electron theory is a better theory of the same particles that Bohr was referring to. I take it that this is the line of thinking that Paul Feyerabend represents. On an account like Shapere's, however, Bohr

would have been referring to electrons when he used the word 'electron', notwithstanding the fact that some of his beliefs about electrons were mistaken, and *we* are referring to those same particles notwithstanding the fact that some of our beliefs—even beliefs included in our scientific 'definition' of the term 'electron'—may very likely turn out to be equally mistaken. This seems right to me. The main technical contribution of this paper will be a sketch of a theory of meaning which supports Shapere's insights.

An 'idealist' theory of meaning, as I am using the term, might go like this (in its simplest form): the meaning of such a sentence as 'electrons exist' is a function of certain *predictions* that can be derived from it (in a pure idealist theory, these would have to be predictions about *sensations*); these predictions are clearly a function of the *theory* in which the sentence occurs; thus 'electrons exist' has no meaning apart from this, that or the other theory, and it has a different meaning in different theories.

The question of 'reference' is a harder one for an idealist: the essence of idealism is to view scientific theories and concepts as instruments for predicting sensations and not as representatives of real things and magnitudes. But a sophisticated idealist is likely to say that the question of reference is 'trivial':† if one has a scientific language  $L$  containing the term 'electron', then one can certainly construct a metalanguage  $ML$  over it *à la* Tarski, and define 'reference' in such a way that "electron" refers to electrons' is a trivial theorem. But if different scientific theories  $T_1$  and  $T_2$  are associated with different formal languages  $L_1$  and  $L_2$  (as they must be if the words have different meanings in  $T_1$  and  $T_2$ ), then they will be associated with different *meta-languages*  $ML_1$  and  $ML_2$ . In  $ML_1$  we can say "electron" refers to electrons', meaning that 'electron' in the sense of  $T_1$  refers to elec-

\*See, for example, the discussion by Hempel, in *Aspects of Scientific Explanation*, (Free Press, New York), (1965), pp. 217–18. A contrasting view is sketched in chapter 13, volume I of these papers.

\*First published in G. Pearce and P. Maynard (eds.), *Conceptual Change* (Dordrecht-Reidel 1973) pp. 199–221.

Reprinted by permission from Hilary Putnam, *Mind, Language and Reality: Philosophical Papers, Vol. 2*. Copyright ©1975, Cambridge University Press.

†In a letter written to Conrad Schmidt in 1895; cf. Marx (1942), pp. 527–30. My agreement is with Engels' realism, not his 'dialectical materialism'.