



## From Vicious Circle to Infinite Regress, and Back Again

Bas C. van Fraassen

*PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1992, Volume Two: Symposia and Invited Papers. (1992), pp. 6-29.

Stable URL:

<http://links.jstor.org/sici?sici=0270-8647%281992%291992%3C6%3AFVCTIR%3E2.0.CO%3B2-C>

*PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* is currently published by The University of Chicago Press.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## From Vicious Circle to Infinite Regress, and Back Again<sup>1</sup>

Bas C. van Fraassen

Princeton University

The demise of foundationalism in epistemology was complete by the time of the Second World War: knowledge and rational opinion do not rest on absolutely secure, self-authenticating foundations, neither in experience nor elsewhere. This realization came to philosophers in large measure at the hands of that same detested logical positivism so often been depicted as foundationalism's last gasp. (Cf. Reichenbach (1938), Ch. 3; in a larger historical perspective, the demise may possibly be dated much earlier.) I will not argue for this; I take the demise for granted. The task which lay, and still lies, before us is to find a way of life after foundationalism. The simultaneous rise of scientific realism *and* a more historical orientation through the work of Hanson, Sellars, Feyerabend, and Kuhn brought this task to awareness. At the same time, it seems to me, these writers opened the way for a truly viable, anti-realist, empiricist philosophy of science—a post-foundationalist empiricism, in contrast to those varieties of empiricism that were identified as the last bastion of foundationalism in epistemology.

With the disappearance of foundations, however, and our subsequent vertigo, there was a great danger of rampant, debilitating relativism. Indeed, some of our colleagues seem to have embraced this with both arms, and now rest peacefully in the idea that Neurath's sister-ship of Theseus can't very well have rudder or sail. Philosophers of science have mostly been in the forefront of the resistance to this temptation (witness e.g. Laudan 1990). It is crucial to empiricism to show that it needn't slide down that slippery slope.

In this lecture I shall confront the putative arguments that the demise of foundationalism leads us into debilitating forms of relativism. I shall argue that they are merely Spectres that still haunt us with unfulfillable dreams of foundations and nightmares of loss when foundations are gone.

The Spectres I shall discuss are connected not so much by their surface similarities as by my diagnosis and solution. In each case I shall direct our attention to certain *pragmatic* aspects of science. I conceive of science as providing us with models. Crucial for us now, in philosophy of science and general epistemology, is to appreciate the autonomy of the act of *relating ourselves* to those models if we are to *use* them. This means, to put it briefly, a certain autonomy of applied science vis à vis pure science. It implies the

need to think in terms of the perspective of individual scientists and scientific communities, and not just concentrate on the content of their theories and beliefs.

### 1. Spectre I: loss of objectivity.

I want to begin with a familiar puzzle about maps. If I am lost, and buy a map, that is not enough. Maps normally do not have an arrow labelled “You are here.” But even if the map I get does have that, the problem is really the same: I have to locate where I am with respect to that arrow. (Imagine I found this map lying in the gutter. Imagine instructions about the significance of such labelled arrows: “If you stand in front of a map under condition C, then...”. You still need to supply the indexical premise “I am in front of this map in condition C.”)

The extra information needed to use the map cannot be encoded in *that* map. When I do have that extra information, I can express it by pointing to a spot on the map and saying “I am there”—a self-ascription of location on the map.

This act too can be described and the information that it takes place can be included on a bigger map (with the label “location of vF’s map-reading at time t”). But of course then what I need to use *that* map is still a self-ascription of location with respect to that map. It does not alter the problem that, with this new map, I can self-ascribe a location by the different words “I am vF and it is now t”. An attempt to replace or eliminate these self-ascriptions leads to an infinite regress, using an infinite series of maps. But even given the accuracy of the whole series of maps, the regress does not succeed in eliminating the need for self-ascription. For I will still be lost, unless I can locate myself with respect to at least one of them, and this I can do only by asserting a self-ascription which is not deducible from the accuracy of those maps.

The topic of self-ascription belongs to pragmatics and not to semantics. That is a fancy way to say that what it does cannot be equated with the content of a map or the belief that a certain map “fits” the world.

*The bearing of this puzzle.* This point about maps is a paradigm for the difference between pure and applied science. The body of science, the totality of accepted scientific information, can in principle be written in co-ordinate free, context-independent form. That is possible for pure science, even if it includes the history of the universe or the evolution of biological species on earth. But to apply this body of science in technology, or even to test it or use it to explain something, or add to it through research, the scientist or scientific community must supply something extra, which does not come with that body of science, but serves to locate the user with respect to it.

Let me put this again, somewhat differently, in terms of models. “Model” is a metaphor, whose base is the simply constructed table top model. We use this metaphor when we talk of cosmological models, Hilbert space models, and the like. We could have used the word “map”, and made maps the base of our metaphor equally well.

Suppose now that science gives us a model which putatively represents the world in full detail. Suppose even we believe that this is so. Suppose we regard ourselves as *knowing* that it is so. Then still, before we can go on to use that model, to make predictions and build bridges, we must locate ourselves with respect to that model. So apparently we need to know something *in addition* to what science has given us here. The extra is the self-ascription of location.

But now the first Spectre appears, and tells us that we have a dilemma. Either we say that the self-ascription is a simple, objective statement of fact, or else we say it is

something irreducibly subjective. In the first case, science is inevitably doomed to be objectively *incomplete*. In the second case, we have also admitted a limit to objectivity, we have let subjectivity into science.

*Historical illustration.* Carnap struggled with this in the *Aufbau* (1928), when he tried to think through his structuralist views about science. He begins Part Two of the *Aufbau* with the announcement “we shall maintain and seek to establish the thesis that *science deals only with the description of structural properties of objects*”. This means that in theoretical science, exactly as in mathematics, what is described is described only up to isomorphism. What theoretical science produces is exactly, no more and no less than, *mathematical models* of physical objects and processes. And this, as Carnap spells out at length, is very far removed from what we ordinarily call description. This view of science is of course not solely Carnap’s; it was also presented by Russell, for example in his *Problems of Philosophy*, and later in Sellars (1965).

As Plato said of poetry and art, so Carnap tells us about science: it is at several removes from reality, it proceeds by means of two stages of abstraction. As first step science describes properties and relations (section 10, p. 19), but then it takes a second step:

There is a certain type of relation description which we shall call *structure description*. Unlike relation descriptions, these not only leave the properties of the individual elements of the range unmentioned, they do not even specify the relations themselves which hold between these elements. In a structure description, only the structure of the relation is indicated, i.e. the totality of its formal properties. (section 11, page 21)

The crucial problem then appears in the next section:

Thus, our thesis, namely that scientific statements relate only to structural properties, amounts to the assertion that scientific statements speak only of forms without stating what the elements and the relations of these forms are. Superficially, this seems to be a paradoxical assertion... in empirical science, one ought to know whether one speaks of persons or villages. This is the decisive point: empirical science must be in a position to distinguish these various entities... (section 12, page 23)

The word “Superficially” makes this a grand understatement. The next seven pages are spent unravelling this “superficial” paradox.<sup>2</sup>

This problem is acute for *any* approach to science which characterizes physical theory primarily in terms of mathematical models. The phrase “mathematical model” is almost entirely redundant in this context, because the only alternative is what you might call “tabletop models,” i.e. concrete structures with labels attached. Even allowing for analog as well as digital models, these are so *finite* in so many ways, that they just don’t get very far if we are talking about theoretical science today.<sup>3</sup> The only manageable abstract structures, however, are what mathematics gives us. And *mathematical description is unique at most up to isomorphism*. We can attach labels, but this “attaching” also has as literal sense *only* the existence of some function, which is in turn subject to the same limitations of description.

Presumably this problem is to be met by turning the tables on Kant’s maxim that in anything there is only so much pure science as there is mathematics (continuing, in his *Metaphysical Foundations of Natural Science*, a rationalist theme from Descartes and Leibniz, which was again recaptured in Duhem’s section title “Theoretical

physics is mathematical physics.”) To turn the tables we should grant this about pure science, or *theory*, but deny that it is all there is to science—empirical science, being more than mathematics, must therefore be more than *pure* theory. An empirical theory must single out a specific part of the world, establish reference to *that* part, and say—by way of contingent, substantial claim about the world—that its models fit *that*. Now, how exactly can this be done?

The problem with that question is that its answer has to be scientifically respectable, and so must itself be regarded as also a subject of science. The task of producing that scientifically respectable answer is exactly what Carnap attempts in the next seven pages. Let us see where it leads him.

In section 13 he says there are two means of fixing reference: by *ostensive description* which relies on perception and gestural indication, such as pointing while one says “*That* is Mont Blanc,” and by *definite description* which singles the object out uniquely by listing some of its properties. He admits that it *looks as if* the use of definite descriptions will be successful only if eventually it relies on some ostensive description. But he reacts to that with “However, we shall presently see that, within any object domain, a unique system of definite descriptions is in principle possible, even without the aid of ostensive description” (*ibid.*, pages 24-25). He qualifies this immediately: that this is in principle possible does *not* entail that it is (really) possible! Indeed, he says whether or not it is really possible in any given case cannot be decided a priori. He adds, and this sounds very ominous:

It is of especial importance to consider the possibility of such a system for the totality of all objects of knowledge. Even in this case it is not possible to make an a priori decision. But we shall see later that any intersubjective, rational science presupposes this possibility. (page 25)

In other words, we might be at the point of a transcendental deduction: the point that it is always in principle possible is very weak, but may become reinforced with the point that it is necessary on the presupposition of, or as precondition for, the very possibility of rational, objective, intersubjective science.

I do not think I need to go through his examples and difficulties in sections 14-16, for the general point can be readily appreciated today. If the universe, or the specific domain under consideration, is invariant under certain transformations—i.e. if it exhibits some symmetry, however abstract or abstruse, description will not fix reference uniquely. Given a particular scientific world picture (and Carnap’s is very, very Newtonian or at least modern still; see section 62 with its very dated discussion of a basis consisting of elementary particles and fields) such doubts about uniqueness of description may look ridiculously skeptical. But given more recent scientific developments, we can’t be so irenic. Both in the case of space-time physics, with its surprising isometries—just think of the various “hole arguments” about determinism—and quantum mechanics with its permutation symmetries, we have just the sort of far-reaching, deep-going symmetries that Carnap worried about.

So Carnap tries to eliminate the need for ostension: what if everything has a uniquely identifying description? He vacillates between two ways of insisting on that: (A) that it is so is a presupposition of science (I suppose that this means we have to believe it whether it is true or not); (B) that we should adopt an ontology in which isomorphism implies identity. But unfortunately, neither would remove the need for ostensive description! In either case, science would presumably, if successful, provide us with “maps” (models) with pervasive asymmetries. But we would still only have a map, and have to

locate ourselves with respect to that. Ostension may have been replaced on the theoretical level with definite description, but it returns as soon as we want to make use of theory.

There is only one solution, the very solution that Carnap dreaded: to let subjectivity into science after all. But what exactly are the implications of this solution? First of all, applied science is autonomous: the *conditions of possibility* of applied science include more than knowledge or belief in the theories and models pure science provides. But what is this “more”? Not a mysteriously different sort of fact which cannot be encoded on the map. The scientific story can be complete in the sense of describing all the facts, including that someone does or does not have the “extra” needed for him or her to apply a particular bit of science. It is just that *describing the having of it* is no substitute for *the having!*

Yet this act of locating myself with respect to a map/model *does involve an empirical hypothesis*. For a particular person to make a specific testable prediction in a given situation requires two ingredients:

[1] Belief about adequacy of the map

+ [2] self-location==> empirical prediction

Here [1] admits of more and less, and [2] can be more or less determinate (“We are somewhere here”). Yet the act is still properly called an assertion, it is a linguistic act, and it can be “refuted” (relative to general beliefs about the map, of course). I will go into more detail on this in section 3, where we shall encounter these acts in a different setting.

We will just have to admit a non-pejorative sense of “subjective”. It is true that this solution gives a special role to consciousness in science. But it does so only on the premise that there is applied science, i.e. there is conscious use of science. The solution entails no more about consciousness than is contained in that premise.

## 2. Spectre II: loss of language

Since the act of locating myself with respect to a map or model is a linguistic act, we must wonder what happens when we apply this same analysis to language. (This was, for example, one of Quine’s great projects, to apply the results of philosophy of science to the language in which we write science—and the outcome he reached was ontological relativity. I will not relate the issues to Quine texts but to another Carnap text.) The second Spectre whispers that in some sense we do not know our own language—we do not know, and perhaps there is not even any fact of the matter about, what we are saying and to what we refer.

Here is a second elementary puzzle about maps; I believe it may help us here. What exactly is it to be lost? Imagine me in Taxco, and compare me with a native born citizen of Taxco. As I walk around there, I become lost: at some point I look around me and I don’t know where I am. I ask this citizen to help me, and he laughs; he is never lost in Taxco, no matter where he goes. But after he helps me, I ask him to draw a map of the town and its environment. The map is meticulous, with very accurate proportions, his house at the center and each landmark indicated with its proper height. This map I take with me when I leave, and during the next year I search many maps and atlases in the reference department of our library. I find twenty-seven areas that are exactly as his map depicts, of cities around the world (perhaps also in other galaxies; this is a philosophical fantasy, after all). Using Xerox, scissors, and paste, I produce twenty-seven maps of those twenty-seven areas, comparable to his drawn map. And the next year, when I return, I show him these maps (perhaps even as parts of a bigger map,

so as to relate them to each other), though all labeled in a strange language he cannot understand. I ask him to show me in which of these twenty-seven places he is. By hypothesis, he cannot. So I assert that he is lost—he does not know where he is!

Of course he will not be taken in by this, unless he is a philosopher too. He will simply tell me that he is not lost at all, and if I will give him one of these handsome maps, he will relabel it for himself, and it will be a very useful map of Taxco! Let us call this the *27 maps trick*.

One historical illustration of how maps can replace reality in the philosophical mind belongs to the philosophy of language, where the maps (models, in the sense of the scientist rather than logician<sup>4</sup>) of parts of our language are “artificial (formal) languages”. Carnap’s book *The Logical Syntax of Language* was written in aid of a grand project: to transform philosophy into a logic of science and *thereby* into a part of syntax. By “syntax” is here meant a certain theory, the theory of syntax, which approaches the living, natural language in which science too is formulated, by constructing purely syntactic systems as artificial languages. Following proper scientific procedure, these are offered as models, of ever increasing and increasable sophistication, for natural language. Then eventually *presto!* natural language disappears and is replaced *in toto* by the constructed language system.

The Dutch logician Evert W. Beth (1963) attacked Carnap’s silent convictions as well as overt declarations about the relation between natural language and the logicians’ constructed language systems. Beth’s criticism is conveyed most easily by thinking about this with a semantic point of view. A syntax has many models (in the logicians’ sense: that is, we can interpret it as a language capable of describing many different structures). It is possible to limit this diversity of models by adding to the syntactic system, besides vocabulary and rules of grammar, also axioms and rules of deduction. But there are limits to this, most remarkably spelled out by Gödel’s incompleteness theorem.

Beth dramatized this with the fiction of another philosopher, Carnap\* who reads *The Logical Syntax of Language*. He keeps agreeing with Carnap for a long time: he admits the syntax as an adequate if partial model of natural language, and then takes Carnap’s axioms and rules as correctly formalizing our own logic, arithmetic, set theory, and so forth. Suddenly there is a disagreement! Carnap exhibits the Gödel sentence which is true if and only if formalized arithmetic is consistent, and Carnap\* says: *but that sentence is false!* Obviously Carnap is very startled by this reaction, since he himself is convinced that it is true, though not provable from the axioms.

How is this possible? Well, both that Gödel sentence and its negation are consistently addable to arithmetic. Hence there are at least two models which could guide one’s intuitions as to truth and falsity, as expressed in natural language. The one model (or sort of model) must be guiding Carnap’s intuitions and the other one (or other sort) guides Carnap\*.

The way in which I described this is however a bit simplistic. It assumes that although natural language does not coincide with a syntactic system, it does coincide with a syntax supplemented with a model or model-type. From the point of view of formal semantics, that is how it has to be described. But that, of course, simply revamps the grand project, casting formal semantics in the role of syntax.

Unfortunately, the same line of argument will demolish that second grand project. For suppose that in naive set theory, we describe a syntax plus axioms *and* its class of models. If we assume this to be an adequate model for the very language in which we work on this

project of construction, then we see a formal replica of our set theory in that model. But again we can imagine a Carnap and Carnap\*. They agree with each other perfectly, it seems, for each agrees that the axioms of set theory, as there formalized, are true. But then, as they move on, disagreement suddenly appears. We have the same impasse as before, though this time it comes from the theorem of Löwenheim-Skolem rather than Gödel's theorem (so perhaps this Carnap\* should have been named Putnam\*).

*Our* situation was even worse than that of Carnap and Carnap\*. For we were listening to both, and agreeing with both, in the conviction that we understood what they were saying—we were even agreeing with both of them. We were convinced they were saying the same thing *exactly because* we were convinced that we understood and agreed with their joint assertions, and *surely* that means we picked out a unique meaning for their words. Then suddenly, disaster, the ground falls away from under us, all along they understood the matter very differently, in two ways. So what was *our* understanding of it then? They turn out to speak different languages which sound the same—so, in what language were we hearing them? What was our language?

The logical character of the argument makes it very general. In fact there aren't just Carnap and Carnap\*, but Carnap\*(1),..., Carnap\*(1000),..., Carnap\*(omega+1),..., etc. Of all these possible languages that sound the same, which one is my language? There is no way to tell, and even worse, there seems to be no fact of the matter.

To diagnose this confusion and to get us out of it, we need to restore a more robust sense of reality. The first thing to insist on is that if someone is talking about such a family of possible languages, the cash value is that he is talking in effect about a model—a mathematical manifold of formal languages which of course themselves are only mathematical entities. The formal languages are models of language games, but the only *real* language games are those which are actually played (in our past, present, or future).

What are these formal languages models of? Here we must distinguish Language, in the sense of the resources we have for constructing and playing language games, from the real language games that are actually played. The formal languages are models of language games (and not a model of the resources; see further my (1986)).

There is certainly underdetermination by the evidence, and perhaps by all the facts, for hypotheses that relate these formal languages, *qua* models, to our language. The only real language games are those actually played, in the past, present, and future. That is quite limited, from a mathematical point of view. Indeed, each of them is presumably only played for a finite amount of time. So Quine is quite right when he points out that there must be, among all the mathematical functions there are, very many *adequate* translation manuals that relate our (actualized) language in different ways to various formal languages. Adequate here means empirically adequate, i.e. fitting all actual linguistic phenomena.

But suppose we make no claims and form no beliefs that go beyond assertions of empirical adequacy for such hypotheses. I certainly do not see that as an epistemic defect. Certainly it means that we leave open many questions about our language, but it does not mean at all that we do not know or do not understand our language. (The criterion for knowledge or understanding had better not be an impossible one to meet!) The many translation manuals are like the many maps, in the 27 map trick, which simply does not establish that the native Taxco inhabitant is lost when he walks around Taxco.

Talk of all possible language games can amount to talk only about all the elements of a model, a manifold of formal languages. This is quite analogous to the case of spatial



perspective: if we want to talk about all possible spatial perspectives, we can't do any better than, or indeed anything really different from, talking about all the points in a certain mathematical space—i.e. talk about a model. In fact, any attempt to talk about all possible language games in a way that goes beyond this, leads into semantic paradoxes.

There are indeed further points to be made here, specifically that *having* a language [or a certain language game being *my* language] cannot be reduced to having “objective” knowledge or belief. But I won't continue with this here—the Spectre that we don't know what language we are speaking is disarmed, I think, once we distinguish the reality of having a language, living in a language [which is like living in Taxco] from the multiplicity of “maps” in the 27 map trick as applied to language.

### 3. Spectre III: loss of experience

I discussed these first two Spectres for their own sake, but also as a prolegomenon to the third, which I find much more threatening. It is the Spectre of what is popularly called Kuhn-Feyerabend relativism. (I am not implying here that either Feyerabend or Kuhn does, or ever did, subscribe to it; the name is quite common and refers to inspiration, which bloweth where it listeth.) The relevance for me is that this relativism implies a “loss of experience” in that it dissolves the observation/theory dichotomy.

When I relate myself to maps and formal languages, my assertions in effect describe where I am and what language is *my* language, the language game I speak. But this language is historically conditioned, it has taken on a certain structure through a historical development, and most specifically through the development of scientific (and other) theories. If I now use this language to report on my experience, those reports are structured and conditioned by the “shape” of my language. It cannot be otherwise, for this is the only language I have. So how can experience be the objective touchstone for science, and how can science have any pretensions to be different from metaphysics? Doesn't the theory-infection of language defeat any demarcation between the theoretical and the observable?

Many people, who consider themselves intellectual offspring of Hanson, Sellars, Kuhn, and Feyerabend, seem to have concluded that once foundationalism was given up, experience *did* turn out not to be an objective touchstone for science, and the development of science obeys no stricter criteria than does metaphysics.

#### 3.1 The third point about maps

Before I explain what this third Spectre threatens, I want to make a third point about maps. When I asked you to imagine that I was lost, and needed to locate myself with respect to a map in order to go on, I am sure you were thinking about an accurate map. Suppose the map is defective or inaccurate; does that pre-empt the act of self-location on it? The fact is of course that all the maps we use leave something to be desired, and some of the most useful are highly distorted. Just think of subway maps: they are inaccurate with respect to shape and distance, and accurate only about some gross topological features of the situation. Yet we unhesitatingly locate ourselves on them, and find that a useful thing to do. Think of a map of the Ptolemaic system of the Universe. Despite our certainty that it is not even topologically correct, there is a right and a wrong way to locate ourselves on it. In one edition of *Paradise Lost* I saw an illustrative map of its Universe, with the Earth at the centre, Heaven above, and Hell below. Now I may quarrel with this, either because it depicts regions definitely not to be found in space-time, or perhaps to say with Marlowe's Mephistopheles: “This is Hell, nor are

we out of it!” Yet I know full well that there is a correct way to locate myself on this map, namely by pointing to the Earth in the middle.

One way in which philosophy of language has always failed philosophy of science is through its disregard of defective language. We communicate in defective language, and know well that we do. We know just as well that we locate ourselves on defective maps, and that the models that are most useful in applied science generally fit real systems only in the loosest ways you can think of. If it is indeed true that past accepted science infects our language in use—and I think that is true, in many ways—then we are always in the position of writing new theories in a language that is highly defective if those new theories are correct. The point applies to maps and language equally: we need, and aim to have, accuracy only in *relevant* respects—inaccuracy elsewhere does not pre-empt the criteria of correctness of self-location with respect to them.

### 3.2 The Scientific Image

Since I hope and try to be an empiricist, I want to resist Kuhn-Feyerabend relativism with all my might. To me at least it is a Spectre, that tells me experience is lost, and indeed that it is lost exactly because the possibility of epistemic foundations was lost. But I think this Spectre, like the others, trades on confusions and misplaced nostalgia.

When I wrote *The Scientific Image*, I followed Sellars and Feyerabend in the pragmatic account of observation. I did this without second thoughts, because I was a student of Sellars. (I did not use the term “pragmatic account”, which was introduced by Feyerabend.) It seemed to me that this account of observation was a central part of scientific realism, so I shared it with my opposition—hence no reason to argue very much for it. I restricted myself mainly to the point that whether or not we can observe something is more or less the same question as whether a person can function as a detector (measurement apparatus) for the presence of that sort of thing (in the sense of measurement in physics; see my (1980, pp. 14-17, 56-59, 80-81)). Of course, this account of observation had as most salient article Feyerabend’s “Thesis I”, nl. the interpretation of observation language is determined by our theories, and changes when those theories change.

I was a little blind to how this might confuse my readers. That thesis was just what gave rise to the sort of relativism that I have just described here as removing the touchstone of experience. Let me therefore now first recount how I see this view of observation, and then elaborate on it.

### 3.3 The pragmatic theory of observation: initial account

Step 1 in the pragmatic theory of observation is the rejection of the Myth of the Given. There are no judgements which are theory-neutral, epistemically secure, and self-authenticating. There are no *foundations* for rational belief.

When my students want to talk about observation reports, I tell them not to take such artificial examples as Russell’s “Red here now!” They should take a typical, representative example, like

(1) Lo, phlogiston escaping!

As Feyerabend indicated, a phlogiston scientist might well shout “Fire”, but he himself, if he had fully believed and assimilated the theory, would regard that as just an abbreviation for (1). Pierre Duhem had already pointed out that in the laboratory, re-

ports take such theory-infected forms, and happily philosophers have rediscovered this several times since then (notably *circa* 1930 and again *circa* 1960).

Step 2 is to point out that we don't learn how to make such observation reports by training in logical inference or translation. Rather we learn by *conditioning* the skill of saying this under certain circumstances, just like Pavlov's dogs learned to salivate when bells rang.

We must be very careful not to misunderstand this step. It looks like another "perceptual mechanism" theory like we find in traditional philosophy (though with a more "scientific" air because of the reference to Pavlov). Instead, this step was the very opposite, a dismissal of such theories, an insistence that they are not epistemology but pseudoscience. (You may recall some of them: the mind grasps the property, then it grasps the substance, and finally the nexus; then with its tiny little hands it puts these together to make a proposition... etc.<sup>5</sup>)

No, the reference to conditioning has only one function here: to show that there is no necessary link between the *content* of the observation report and the *conditions* to which it is a response. The point of the metaphysical mechanisms of perception stories was to secure such a link *a priori*. That is a mistake. We could condition a person to yell "Bingo" when he sees fire, or to say "There goes an alpha particle" when he sees a trail in a cloud chamber, or to utter "Another flying saucer!" whenever he sees a light in the sky. By means of this conditioning—whatever that process may be, we don't care—a correlation is established that makes the person in effect a reliable measuring instrument or detector of the conditions to which he is conditioned to react in that way.

Feyerabend called this account "pragmatic" because of the role he assigned to interpretation when he elaborated on these first two points. One example he seems to have had in mind goes like this. Let us assume that at one time the phlogiston theory had the kind of cultural hegemony that Lavoisier's theory of oxydation acquired afterward, and has now. In the presence of fire, people cried "Fire!" both then and now. In the earlier period, they and their contemporaries took that to be in effect short for "I am in the presence of phlogiston escape." We say that then too, people were, with those shouts, reporting the presence of fire—by which we mean, of course, of oxydation.

There is a difficulty with this. It is certainly true (you and I agree) that people typically shouted "Fire!" when in the presence of oxydation. But when they did so, it did not seem to them *that* they were in the presence of oxydation, *but that* they were in the presence of phlogiston escaping—and that is what they meant with their words. There is something Pickwickian, therefore, in the *oratio obliqua* report of their utterance as: they reported, reliably, that oxydation was occurring when they were in the presence of this phenomenon.

It may be that Feyerabend also had in mind another type of example, in which perhaps we know only that in a certain medieval culture people reliably shouted "Vuur" when they saw fire, and we have lost all knowledge of their language and beliefs. We have here a problem of radical translation, and perhaps Feyerabend thought that the correct, or at least an admirable, translation of their shout "Vuur", by us, is "I am in the presence of oxydation." But this has the same oddity, for in fact (surely?) in our opinion they most likely did not mean that. For to say *that*, they would have had the concept of oxydation, hence something like Lavoisier's theory; and it seems unlikely that they did.

Blindness and insight: the insistence on the historical character of language, and its infection by theories, with the consequent "shaping of experience"—the very move

that gave salience to the roles of history and interpretation—was here apparently accompanied by a certain ruthlessness with respect to interpretation, a readiness to just impose interpretation to the point of anachronism. We need a more careful elaboration of the notion of “observation report.”

### 3.4 The pragmatic theory of observation: elaboration

The term “observation report” is a technical term, a bit of philosophical jargon. Explication must respect the purpose for which it was introduced, and fit the paradigm examples that it was meant to fit. We need not, of course, respect all the beliefs philosophers may have had about those examples. The most important of these paradigm examples are certain utterances in the laboratory or observatory. The reports issued by such institutions convey the data already reduced, summarized and corrected by statistical methods. But those summary reports are based on individual reports by trained observers—and *those* are the paradigm for observation reports. Undoubtedly these are in focus because empiricist philosophers see there the ultimate touchstone, the bottom line, among the criteria of success in science. Other examples classed with these involve shouts of “Fire!” when people see fire, and “Red here now” if they have been specially hired to inspect color samples or look for blood stains or are students of Bertrand Russell. With all this in mind, let us try to explicate this technical term anew.

Under what conditions do we classify an utterance as an observation report? I raise the question in this way, to allow full play for the role of our own beliefs, which is at least one thing meant when people say that we “interpret”. For of course we classify on the basis of our own beliefs. (Consider the alternative!) This is innocuous, for it applies equally to our classifying of butterflies and minerals.

In answering the question, we must distinguish between very strict and looser or derivative uses of “observation report”. We should also separate cases in which we already understand the words used, and when we do not. (For the sake of example, I assume that we today do still understand “Phlogiston is escaping”, though the phlogiston theory has long since been rejected.) The central case first: the strict sense, for the case when we do understand the word. Assume that people X reliably (with high correlation) utter ‘E’ under certain conditions  $f('E')$ . One example of this, which I take it is not an observation report, is that people reliably cry “Mother!” in situations of mortal danger.

Special conditions for the central case:

- (1) The utterance of ‘E’ is symptomatic of its own truth
- (2) On the occasion of utterances of ‘E’, it seems to the utterer that E
- (3) ‘E’ is a report on the occurrence of an observable event, process, or state of affairs (which includes the presence of an observable object)
- (4) The utterance of ‘E’ plays the role of a self-locating self-ascription, for the utterer, with respect to some map, picture, model, or general description of the world. (That is what is right about Russell’s “Red here now”: observation reports are indexical, acts of self-location.)

I think that (3) will raise the most eyebrows, but let me comment on the whole set of conditions first, and then on each condition separately.

These conditions make no sense at all for cases in which we don’t already understand ‘E’. Moreover, they are very strict. In the very strict, “central” sense of “obser-

vation report” which they mark out, I (Bas van Fraassen, in 1992) do not classify “Lo, phlogiston escaping!” as an observation report. The reason is that I do not believe its utterances to be symptomatic of their own truth, since I do not believe that phlogiston exists. For the sake of the example I can easily imagine a chemist before Lavoisier who did so classify it, and was right to do so (because his beliefs were different from mine). But of course, this strictest sense is not the sense in which we always, or even typically, use the term. I call it central because I think that all our uses are derivative from it in obvious ways. It is at best a simple exercise to construct the derivative senses which weaken these conditions, so as to sanction such derivative uses as are exemplified in:

- (a) Peter’s observation reports are reliable but couched in language infected by a false theory
- (b) Peter’s unintelligible cries are in effect observation reports of type...
- (c) Peter’s observation reports are intelligible and infected only by theories we also believe, but extremely unreliable.
- (d) Peter’s “Fire!” utterances are observation reports about fire, but only when not inebriated, which is rare in his case.

“Phlogiston escaping!” as said in the phlogiston theory’s heyday should probably be described by us today as an example of type (a). I will leave the formulation of these derivative senses and examples as an exercise.<sup>6</sup>

In (1) the condition-content link is reinstated, but note well: not as *a priori*, “internal”, “broadly logical” or whatever, but in the form of a factual condition of high correlation (implied by our beliefs about the utterer’s performance). Clearly we can consider this condition met for highly theoretical assertions, but only when we believe sufficiently much of the theory.

The use here of “symptom” is of course post-foundationalist; we do not require certainty, only high correlation. But this too needs to be carefully formulated. We hope for instance that the conditional probability (X is in the presence of fire *given that* X shouts “Fire!”) is high. We do not add that there is a high probability of his shouting “Fire” when he is in the presence of fire—he may often find it unremarkable. But we do want to add that the conditional probability (X shouts “Fire!” *given that* X is not in the presence of fire) is low—the symptom must separate the condition from its rivals. More detailed discussion of this sort of requirement should take its cue from the detailed treatments of the general notion of measurement in the foundations of quantum mechanics.

Point (2) requires that the person uttering ‘E’ is not merely parroting, but really reporting on how it seems to him at that time. There is an important less strict, derivative sense not included in the examples above. For the sake of efficiency as a participant in research, an investigator may thoroughly immerse himself in the world picture of accepted background theories—suspending his disbelief or agnosticism with respect to some of their implications. This person is then not only speaking but thinking and readily responding *under supposition* of those background theories. I take this to be a common and indeed prevalent phenomenon, not only in scientific research but in virtually all types of civilized dialogue.

Point (3) makes no sense unless the question, whether the events described by ‘E’ are observable, is independent of the question whether its utterance is an observation report. Since typical examples of ‘E’ are sentences in highly theory-infected lan-

guage, this requires that the observable/unobservable distinction is theory-independent. This I will take up in the next subsection.

Point (4) I consider the most important, besides (1). For the importance of observation reports derives very largely from their being (in Sellars' terms) "language-entry moves." The first three conditions can be met by examples like "It rains more than three times a year in Indianapolis" or "It either rains or doesn't rain."

Our entire discussion of self-locating self-ascriptions was *a fortiori* a discussion of observation reports. They have as complete text an indexical assertion that locates the speaker, on that occasion, in some definite part of his own general, "objective", world picture. I realize that I am leaving this somewhat metaphorical or analogical: it is not literally true that we carry our "general" opinion with us in the form of a representation that encodes it, like a picture or map. But we can reasonably represent ourselves as doing that, for present purposes.

Finally, let us consider in what sense the *status of observation report* is "relative" or theory-dependent. As far as we can tell at this point (pending discussion of condition (3)) it is not. Imagine e.g. that Paul believes that

(5) It seems to Peter that phlogiston is escaping.

The beliefs of Paul—who may here be classifying Peter's utterances as observation reports—do not enter into the truth conditions of (5). Whether or not (5) is true does not depend on Paul at all. Of course (5) is not true unless Peter has the concept of phlogiston, so Peter's mental life does enter into those truth conditions—not surprisingly, since (5) is about Peter.

There is no non-trivial sense in which (5) is theory-dependent or theory relative. This point is in no way affected by the fact that Paul cannot believe (5)—and hence may not be able to classify a certain utterance as observation report—unless Paul has the concept of phlogiston. (Think again of classifying butterflies or minerals.)

Yet, given the possible differences in beliefs and concepts between Paul (the classifier) and Peter (the utterer), it is much more fruitful to focus on classification by a particular person (scientists, scientific community), than to consider this topic in the abstract. There are of course philosophical positions according to which there is no fact of the matter as to whether it seemed to Peter that phlogiston was escaping (or whether he meant phlogiston when he said "phlogiston", or, I suppose, whether he refers to himself when he says "I"). I am only concerned to point out that nothing in the pragmatic theory of observation so far supports such positions, though of course each could in some way be grafted onto the other.

### 3.5 The empirical character of observation

The term "observation report" is, as I said, a technical term of philosophy. The word "observe", on the other hand, has a common use, more or less the same as that of "perceive." The same is true of the word "observable" (though it has also been given technical uses, for example in quantum mechanics). In philosophical discussion I take it that it is meant to have its common use, unless otherwise indicated.

This term "observable" is very much like such other common words as "portable" and "fragile". They are, so to speak, anthropocentric terms, for they refer to our limitations. They are not person-centric, however; laptop computers are portable and

wine glasses fragile, even though some people are too weak to carry or break either. The limitations they refer to are of the human organism. So far, not even philosophers have suggested that the demarcation of the fragile has shifted after the development of such sophisticated instruments as the sledge-hammer. I want to go further, however, and also insist (as I did in *The Scientific Image*) that the observable-unobservable distinction is in no important sense theory-relative or theory-dependent.

To focus our questions, consider the following statements:

- (1) X's utterance of "Phlogiston escaping!" was an observation report
- (2) X observed that phlogiston was escaping
- (3) X observed a phlogiston escape

None of these can be equated with any of the others. Moreover, (1) does not imply (2) or (3); and (3) also does not imply (2). An observation report is only symptomatic of its own truth, and symptoms don't guarantee more than high probability. Therefore, if the observation report was issued, it does not follow that the reported phenomenon was indeed observed.

There is good reason to keep this terminological break. We want to cite the observation report as *evidence* that a certain phenomenon occurred. That means that we need to be quite confident that there were such observation reports, before we have established our conclusion that there were any such phenomena. But (2) does not add much to (1), if anything, beyond the endorsement of (1) as veridical. In addition, if (1) and (2) are both true, then certainly (3) is true.

But is there some sort of theory-dependence in any or all of them? Statement (2) cannot be true unless X could on that occasion give us the observation report "Phlogiston escaping!" This requires that it seems to him that phlogiston is escaping, which requires a belief in phlogiston, and even more, the concept of phlogiston.

But (3) is quite different. You and I do not believe in phlogiston. But we can suppose for a moment that the phlogiston theory is true. On that supposition, all those fires we see are phlogiston escapes, and therefore, still on that supposition, we see phlogiston escapes. This supposition does not at all entail however that we believe in the phlogiston theory, or even that we have the concept of phlogiston. Therefore, whether or not (3) is true is quite independent of what theories or concepts X has.

Let me restate this crucial argument in another way as well. From my present, post-Lavoisier point of view, the following three premises are jointly satisfiable: the *first premise* is the phlogiston theory, the *second* is that people see fires, and the *third* is that no one has heard of or understands the phlogiston theory. Only the second premise is true, but my conclusion follows from the fact that they are jointly satisfiable. (My argument presupposes that "X sees Y" is extensional; this is the conclusion of my Stone Age Native example (1980, p. 15).) If all three premises are true, then people who have no concept of phlogiston are observing phlogiston escapes. Since the three premises are jointly satisfiable, it follows that what is observable is not theory-relative or theory-dependent in any important sense.

It would therefore be quite wrong to try and explain (3) in terms of what observation reports or perceptual judgements people can reach. It is only under certain special conditions that the two are connected, and those conditions are merely contingent facts. This really follows from the absence of any possible *a priori* condition-content link.

The demarcation of what is observable and unobservable concerns not (2) but (3). The question of the conditions under which such statements as (3) are true is not a philosophical question but an empirical one. If anyone wants to frame opinions about just what is observable, I would urge him to draw on physiology and psychology, and empirical science in general, and not to ask philosophers at all.

Of course, we should expect then that *our opinions about what is observable* will change as science changes. But that does not mean that what is observable changes too. The recent paper by Laudan and Leplin (1991) falls into exactly that confusion. They point out quite correctly that what we regard as observable is not constant across the history of science. But then they conclude that the line between what is and is not observable has shifted right along with it.

Well, our opinion of the amount of water present on Mars is also not constant across the history of science. Yet the mass of water on Mars has not been shifting along with this shift in opinion. Our judgements of empirical adequacy of theories will of course vary; but whether those theories are empirically adequate—just like whether or not they are true—is a characteristic which they do not lose when we begin to think differently.

### 3.6 A query about epistemic policy

And yet the Spectre of so-called Kuhn-Feyerabend relativism has not disappeared. For after all, the only input for our epistemic policies or decisions will be our opinions, and not the facts themselves. Probably the most contentious point in *The Scientific Image* was my submission that science allows us to not believe the theories we accept, but to believe only that they are empirically adequate. Science, according to me, does not require more belief than that (though of course you may believe more if you wish, for whatever reasons of private satisfaction you can muster).

This distinction rests on the theory-independence of empirical adequacy, i.e. of the observable/unobservable distinction. But when someone tries to implement that policy, his input is of course not the separation in nature of observable and unobservable objects and processes, but his *opinion about* that demarcation. This opinion is historically conditioned by the language in which he frames all his opinions, including his observation reports—and that language is infected with theories he believes, and which may well be false. So in practice, the pursuit and correct realization of this empiricist epistemic policy will be heavily perspectival, and depend on the historical character of the language of the scientist or scientific community in question.

But after the distinctions I have made, this can hardly sound like a debilitating relativism. It is true that when I am forming and changing my opinion, I factor in what I classify as my own observation reports—and these are not epistemically secure foundations, they are full of empirical risk. Well, what would be the alternative? If we don't give in to a nostalgia for foundations, there is no alternative. And if there is no alternative, we can't very well be said to have any defect in our epistemic life because of that.

We should distinguish here between two cases. When a new theory comes along, and we can draw a distinction for ourselves between its being true and its being empirically adequate, then we have a choice. We can then follow the policy of accepting without believing. That is the first case.

But the structure of our language may be such that we cannot make this distinction, and therefore have no such choice. That is the second case. This means that, as we see it at that time, the new theory is only about what is observable and does not at-



tribute unobservable structure to the world. The supposed problem is now that this judgement on our part derives from theories which infect our language, and thereby our opinions about what is observable—and those infecting theories could be wrong.

If we genuinely have no choice here (between mere acceptance and acceptance with full belief), then those infecting theories have no name, and we are incapable of conceiving of anything contrary to them. No standard example of a scientific paradigm ever achieved that status. Newton's physics reigned for 200 years, and didn't achieve it. We are now 500 years beyond the Renaissance revival of the atomic theory, and it still has not achieved that status. But still the Spectre says that there may be such infecting theories, and they may be false, or at least they may be contrary to theories that infect the language of other historical communities in like manner.

Well, the Spectre of Kuhn-Feyerabend relativism has now been reduced to this worry, that I am hampered or conditioned (in some pejorative sense) by possibilities of which I cannot conceive. Only Cartesian foundations could save me from such a fate, and no such foundations are possible. Any standard of rationality that makes this a defect is a standard no human could meet. I conclude that this Spectre too is only part of the vertigo that overtook us when we finally let go of the search for absolutely secure foundations.

#### 4. Spectre III: loss of value

The three Spectres confronted so far were illusions of loss: illusions that, having lost epistemic foundations and being lost among paradoxes, we had lost all. That is not so: in rejecting foundationalism, we have rid ourselves only of unsustainable illusions. Loss of an illusion is not a loss, though it give rise to an illusion of loss. But my arguments designed to disarm those Spectres had a presupposition which must have become more and more obvious as I went on. That is the presupposition that we mariners at sea, who are building and rebuilding our ship of scientific opinion about the world, can rationally evaluate what we are doing.

##### 4.1 The insinuations of naturalism

If rational evaluation—that is, rational value judgment—is impossible then the enterprise of epistemology is vacuous after all. The fourth Spectre tells us that evaluation makes no sense without a *basis* for evaluation, a foundation of value. But by now there is little hope for foundations of any sort.

We may associate such phrases as “the death of epistemology” more with Richard Rorty than with Quine, but Quine preached them more effectively. In “Epistemology naturalized” he presents a quick sketch of orthodox, foundationalist empiricism as developed and also destroyed *malgré eux* by Hume and Carnap. Then he calls *that* epistemology and pronounces it a failure. Quine concludes that epistemology was an enterprise mistaken in intent, and that the only genuine questions left over from it are questions of fact, which fall within the scope of empirical science itself.

What exactly does Quine, and those who follow him here, advocate? In one way, epistemology had already been thoroughly naturalized exactly when the pragmatic account of observation replaced the metaphysical mechanism accounts thereof. The whole psychologistic tangle of the Given and its transcendental processing had been swept away. The evaluation of scientific opinion about the world we live in (*internally*: how simple, unified, coherent, informative? *and externally*: how reliable, how well calibrated, how good a fit to the phenomena?) and of scientific methodologies

was free to proceed without its previous encumbrances. So what was the *extra* naturalization which Quine advocated?

Epistemology had traditionally concerned itself with the evaluation of belief and opinion: is it knowledge, is it rational, is it reliable? Quine concludes in effect that the very possibility of such evaluation is a mistake; he concludes that the only significant core of such putative value judgments is a certain factual component, within the scope of (cognitive) science itself. One suspects (and this is supported by the very little Quine wrote specifically on the topic) that this contention is part of a certain thesis about value in general. That thesis about value is not new; it was already announced by the logical positivists a long time ago. A.J. Ayer brought the message to England sixty-odd years ago:

in so far as statements of value are significant, they are ordinary “scientific” statements, and in so far as they are not scientific, they are not in the literal sense significant. (1946, pp. 102-3)

Among these statements of value Ayer included evaluation of epistemic rationality, as I shall discuss below. Today this message comes in the voice of naturalized epistemology and the cognitive or naturalizing turn in philosophy of science.

The conclusion that science is a natural phenomenon within the scope of one or other branch of science is of course true and innocuous. What is not so innocuous, I think, something that this conclusion insinuates. The insinuation is that upon proper disinfection and sterilization of philosophical discussion, all questions concerning value are systematically eliminated in favor of factual questions, from which evaluative terms are really absent.

“Naturalism” is an accordion term, like for instance “realism”. It is illuminating to look at Wilfrid Sellars’ contemporaneous development of an (in many ways) equally naturalistic philosophy. For unlike Quine, Sellars did not eliminate value judgments. Sellars’ version of eliminative materialism envisaged a final language of physics whose resources included not only description but also the expression of intentions. That entails means for expressing value judgements, because his specific theory of values explains them in terms of intentions. To have applied science—more generally, to live and act in the world—we cannot do with factual, descriptive language alone. Since life includes applied as well as theoretical science, the language in which it is carried on must be expressive as well as descriptive, and therefore elimination of the category of value judgements from our thinking and discourse is impossible.

Quine’s and Sellars’ philosophies can both claim to have left epistemological foundationalism behind. They agree, we may take it, on the irreducibility of value to fact. But one view dismisses and eliminates, while the other gives an important if untraditional place to the category of value judgement.

#### 4.2 Instrumental rationality (1): elimination of norms?

In what sense, and to what extent can evaluative judgements be replaced by factual judgements? An example would be the evaluation of some methodology, which consists in a policy for revising opinion [including belief in theories or hypotheses] under a range of contemplated, possible circumstances. The evaluation might be *comparative*, for example, a judgement about whether Bayesian methodology is better or worse for sociology or psychic research than orthodox statistical testing. Or it might be *qualitative*, yielding only the judgement that a given methodology is minimally satisfactory, or at least rational, to adopt in a certain branch of science.

I used the code-word “rational” deliberately, knowing that it is like a red cape to a bull, because of its associations with *a prioricity* and rationalism. As Ron Giere (1985, 1988), among others, has effectively pointed out, however, there are concepts of rationality which have no truck at all with *a priori* or absolute standards. He uses the term *instrumental rationality* to refer to evaluations of effectiveness of means with respect to given goals, under given circumstances that include a specification of all relevant resources.<sup>7</sup> A judgement of instrumental rationality is purely evaluative, for it says, for example, that methodology X is rational. But whether the judgement is correct, appears to depend on a certain factual question: namely, whether X is an effective means to certain ends under certain circumstances. It appears then that by adopting the notion of instrumental rationality as the evaluative standard, we have effectively eliminated evaluation as such in favor of factual judgement.

*But notice two things.* The first is that exactly the same happens if instead of instrumental rationality we adopt one or other rival conception of rationality. Suppose for instance that *a priori* rationality consists in being in accord with certain criteria which are claimed to be *a priori* norms of practical reason or the like. Then the judgement that methodology X has this feature of *a priori* rationality is also correct if and only if a certain factual statement is correct (namely, that X satisfies those criteria). This apparent elimination of the normative is not due to the peculiarities of the instrumental concept of rationality. It is due simply to the logical point that *whether or not an evaluative judgement is correct is itself a purely factual question*, once a precise standard of correctness has been specified.

The second thing to be noticed is that the consideration of instrumental rationality, or any other sort of evaluative category, *loses its point entirely unless* there is also intellectual activity which does not consist purely in factual judgement. The ends with respect to which methodology X may or may not be effective must be adopted, *must be someone's ends*, or the question is moot. Of course, we as philosophers, disinterested super-human critics that we are, may be above all that. We may describe a range of possible ends, and ask about the effectiveness of X with respect to each. But even then, for our activity to be what it purports to be, we must have adopted some ends ourselves. Otherwise there is no such thing as correct/incorrect for our own statements. As beautiful sounds and writing they may fall short, as giving us private pleasure they may be a great success; but their relevant success will be defined by our goals. Unless certain cognitive ends are *our* ends, our statements are mere phonetic display. (This is part of a crucial more general point in (Putnam 1982)).

Value judgements too are self-locating self-ascriptions. They differ from the “I am in Taxco” type in that they locate the judge not (only) with respect to a map or model of the factual landscape but with respect to a tabulation of value scales or standards: “(this is where I am, and) that is my/our standard among those which can apply to this place”.<sup>8</sup> In addition, their linguistic role is not simply to make autobiographical statements of fact, but to affirm or express the evaluative propositional attitude, or the commitment to those values.

In one sense therefore normative or evaluative questions are *always* automatically eliminated in favor of factual questions as soon as they are made clear—and in another sense, normative or evaluative questions *never* disappear, are never eliminated. This is so exactly because science is applied science as well as theoretical science, and life is practical life as well as cognition, even in our most intellectual moments. Naturalizing reason leans on a banal truth to insinuate a false conclusion.

### 4.3 Instrumental rationality (2): relativism?

Focusing on instrumental rationality may perhaps be no more than giving the idea of rationality its proper due; in any case (as I have just argued) it certainly does not remove us from the realm of value. But if value is not lost altogether, is it made relative? So relative as to become empty?

When A.J. Ayer brought the logical positivist gospel to England, he made one very prudent rhetorical move. He wanted to give pride of place to scientific ideals of rationality, among all values; but he also wanted to promulgate the emotive theory of value. What he did very prudently was to present the special case of rationality first, and the general one of value afterward. In that way he could sound commendably *naturalistic* about scientific rationality and dismissively *relativist* about value in general. His account of rationality was that of instrumental rationality as gauged by the person's or community's own lights:

to be rational is simply to employ a self-consistent accredited procedure in the formation of all one's beliefs..... For we define a rational belief as one which is arrived at by the methods which we now consider reliable.... And here we may repeat that the rationality of a belief is defined, not by reference to any absolute standard, but by reference to our own actual practice. (1946, pp. 100-101)

We can all agree to this when the “we” and “our” refer to a community which we at the same time *belong to* and *endorse*—the scientific community whose values and ends we *underwrite* as our own. Giere's characterization of instrumental rationality as “using a known effective means to a desired goal”, “policy... based on solid empirical findings about effective strategies for pursuing various scientific goals” (1988, pp. 7 and 10) coincides with Ayer's at least in that special case. But what do these views entail for judgments of rationality that cross community boundaries?

The view of rationality that emerges from my two quotations of Ayer is the following. The judgement that a person, policy, action, or belief is rational is an evaluation and therefore has two parts (aspects, sides). It expresses the judger's attitude, his “ranking” of the object in question—to this extent the judgement is not a statement of fact, indeed not a statement at all, and “not in the literal sense significant.” But it also expresses a factual assertion, to the effect that a certain procedure is e.g. self-consistent, accredited, reliable. The connection between the two is this: (1) such terms as “accredited” or “reliable” are elliptical, and tacitly refer to certain standards; (2) when these standards are *our* standards, our evaluation (“ranking”, evaluating attitude) is based exactly on the extent to which the object meets those standards.

The simple case is the one in which actor and judge belong to the same relevant community, sharing the same relevant standards. In that case, if Peter says that Paul is rational, and explains that Paul uses a self-consistent, accredited, reliable procedure, there is still an ambiguity: consistent, accredited, reliable by whose standards? Peter's or Paul's? But the ambiguity does not matter, for the standards are shared. We can discern three possible views. The first, which is arguably Ayer's own, is that the judge ignores the actor's standards, and makes tacit reference only to his or her own. If I say that the Romans built bridges *rationally* then I mean, *pace Ayer*, that they used “methods which we now consider reliable.” I doubt that this is a correct construal, for it does not account for the difference between being irrational and being factually mistaken about means and resources.<sup>9</sup>

The second view perhaps no one holds, so let us attribute it to the fictional philosopher Ayer\*: that the judge makes reference only to the actor's standards. Hence if I say that the Romans waged war *rationally* then I mean, *pace* Ayer\*, that they did so by procedures accredited and regarded as self-consistent and reliable in their community. This is a concept of rationality as pure coherence. Indeed, it deletes from consideration all factors except the actor's opinion. For what is accredited or regarded as reliable will depend on their own opinions about what their goals are, rather than on what those goals really are (if there is a difference). Similarly it will depend on their own opinion of what means are available, rather than on the means truly available. The advantage of Ayer\*'s view is that the evaluative aspect of a judgement of rationality is utterly diminished in importance: only the facts about the actor matter at all. Or so it seems.

That is simply not the role of judgements of rationality. In the first place, when I gauge X's rationality, I ignore some of X's own judgements. For example, I ignore his opinion about whether or not he is rational, and even to some extent about what means are effective. I do look at his opinions about what his goals, means, and resources are like, and see if *those* entail that the means are effective. Consider the dialogue

- A. I don't fasten my seat belt, because I believe that my safety does not require that.
- B. What if you had a collision with your safety belt not done up?
- A. That is unlikely, given that my safety does not require fastening my seat belt.

If A cannot marshal opinions about seat belts and accident statistics independent of his first expressed belief, we will not judge him to be rational.

There is a deep-going difficulty for Ayer\*'s construal. Not having a foundationalist epistemology, we cannot think of A's estimates and probabilities as determined uniquely by a basis of evidence uninfected by his own epistemic attitudes. If we look for "basic" beliefs in A's opinion, some of them may be assertions of safety, reliability, and effective means. His probabilities should be coherent with each other, and cannot be "taken apart" even to *that* extent. In this example, I would still judge A to be irrational even if I thought his opinion was coherent. My explanation would consist in *my selecting* certain of his opinions as "his evidence", and evaluating how his opinions go beyond *those* by my own standards for prudent extrapolation and risk assessment.

It may be remarked here that I am pointing to a critique of A's factual opinions rather than of his standards. But that point is moot since, as I remarked, Ayer\*'s construal leaves no room for anything else. If all that matters in the assessment of A's rationality is how his conduct looks by his own lights, then the only remaining target is his *opinion* about his conduct.

The third view is more easily read into Giere's words than into Ayer's, so I shall tentatively attribute it to Giere. Let us refer to a person's goals and opinions together as his *perspective*. Then this view is that a judgement of rationality is an assessment relative to the judge's perspective of a relationship between the actor (or actor's conduct, policy, opinion, beliefs) and the actor's perspective. After the difficulties with Ayer's and Ayer\*'s one-sided views, this view is to be recommended. Spelling out the relationship in question is not easy, however. The evaluative judgement will, on any given occasion, rest on a selection from the judge's goals and opinions on the one side, and on the other side a similar selection (by the judge) from the actor's perspective. The attitude expressed derives from the judge's perspective *plus* his selection [surely one with a good deal of leeway] of opinions and goals of the actor (to which to relate

the actor's *other* opinions or goals or the actor's conduct, etc.) There is a very strong element of subjectivity here; that is, the judge's perspective intrudes crucially in this value judgement. But that is what makes it a value judgement, an evaluation.

If this is correct then reason can indeed not be naturalized, at least not to the extent that questions of value can be effectively removed from epistemology. On this view, assessment of instrumental rationality too is a value judgement. It too is the expression of an attitude which exists at all only because certain standards are *our* standards, the ones we personally or communally adopt or endorse, and which define *our* commitments. Epistemology is then not limited to research into actual scientific or common practice, but must include exploration of which policies (for scientific research, for change of opinion) are rational—a question which is not equatable to any value-neutral, perspectiveless, commitment-free question.

#### 4.4 The last Spectre's last stand

My discussion was limited to instrumental rationality which may or may not be typical of the subject of value. It may be objected however that even if judgements of instrumental rationality are genuine and ineliminable value judgements, the victory is Pyrrhic. The Spectre that whispered of the loss of value may be content to see it thus saved. Our evaluations of opinion, action, and methodology as instrumentally rational may be important for us, for the simple reason that they express our perspective, the stance we take, the goals we have adopted, the standards we set. But theoretically speaking, this may be idiosyncratic: is any such value judgement invariant under all shifts in perspective?

Invariance is crucial to significance—in factual description. The “perspective” metaphor to which I helped myself insinuates, by the force of its analogy, that a similar criterion of significance applies to value judgement. But let's not be the prisoners of our own metaphors—especially not if we borrow them from the philosophical community at large. This particular metaphor is striking in part because of its passivity, its suggestion of the “I am a camera”, seeing-eye image characteristic of the Cartesian ego. While I included goals as well as opinions in one's perspective, the use of this visual metaphor may subtly subvert itself by removing goals from our attention.

What happens in a shift in perspective? Giere rightly points to Kuhn's lesson about the diversity in goals and standards among scientific communities as a difficulty for *a priori* notions of rationality. Partly thanks to Kuhn's *Structure of Scientific Revolutions*, philosophy of science began to admit the following view:

(K)two scientific communities, even if starting with the same scientific background and receiving the same data, may nevertheless reach different conclusions and accept different theories, without thereby showing any defect of reason in either.

Such differing communities will accordingly also differ in their judgements of instrumental rationality. Are we not pushed into admitting the following possible scenario: we judge a certain procedure P to be instrumentally not rational while in community X the same procedure P is judged to be instrumentally rational—but we judge X's evaluation of P as rational as being itself instrumentally rational? If that can happen, does it not show that judgements of instrumental rationality are so feeble, and so self-undermining, that they are of no importance?

Actually this scenario could easily happen already if there is a factual disagreement between our community and community X about the reliability or effectiveness of P

with respect to the same goal. We certainly would not conclude that the issue at stake is unimportant just because we think of X as rational. So the argument couched in the rhetorical questions above has no general validity. But what about the more worrying scenario extended from the above, in which we and X agree on the facts? Then our disagreement must derive from differing standards or goals, from what we think important and they regard as useless or laughable. But in that case, *by hypothesis*, we would not consider the issue at stake as unimportant either! The argument in the rhetorical questions placed us “between” two perspectives, *suggested* that we imaginatively remain neutral between the two, and then asked us to conclude that the value judgements were of no importance. But “of importance” is a value-term, which makes tacit reference to a perspective when used—namely to the user’s. It is logically incoherent to ask us *not* to consider whether an issue is important to us, or to community X, or to a specific other community, *but yet* to answer whether the issue is important. For our answer, if genuine, can only express, therefore reveal, what is important to us.

The idea that there are in principle many different scientific communities, all equally rational, all of them right in their judgements of scientific rationality, and yet disagreeing in those judgements, is subtly incoherent. For the phrase “equally rational” and “right” are both evaluative. Therefore the assertion in question, if genuinely used, would express the user’s attitude, and his judgement would be incoherent. It is a recurrently appealing and befuddling idea. Its core of truth we have seen already, but its fascination derives from the Spectral whisperings surrounding that core. These whispers manage, strangely enough, to make us ask ourselves a question on the order of “Will you keep this for me until I give it back to you?”

There is no view from nowhere. If it looked for a moment as if naturalized philosophy of science had reinstated the objective view from nowhere, by shifting to the concept of instrumental rationality, that was an illusion too.

The Spectre bandies about millions of perspectives conflicting with our own, each of which could have been *our* perspective after different histories that shared a common beginning. In these different perspectives we see different goals and very different opinions about what the world is like. Looked at ‘from above’, which should we endorse? Would it not be totally *arbitrary* to endorse our own, the one we actually have, and say *we live here*, those goals are the ones which are *worthwhile*, that is what the world is like?

But by hypothesis that is the one we endorse! Endorsement reflects our own perspective, and is not endorsement if it doesn’t. To say that we are arbitrary unless our own endorsement is perspective-free is to hold us to a logically impossible standard, asking us to judge without judging. If rationality and objectivity were identified through such a standard of non-arbitrariness, then they could indeed not exist; but we should not hanker after the logically impossible.

## Notes

<sup>1</sup>I want to acknowledge many helpful discussions and correspondence relevant to this paper, especially with Richard Foley, Ronald Giere, Elizabeth Lloyd, Hilary Putnam, Eliot Sober, Ernest Sosa, Fred Suppe, and Nobuharu Tanji.

<sup>2</sup>All quotations are from the translation by R. A. George (1967). See further page 107 and the paradoxical opening of sect. 180 (page 290).

<sup>3</sup>I strongly object to the assimilation of models to analogies and metaphors (as opposed to, perhaps, the converse) which give rise to such characterizations as “a model in science is a systematic analogy postulated between a phenomenon whose laws are already known and the one under investigation.” There are indeed such special cases; for instance, one might say that at a certain historical moment, gas diffusion was a model of heat diffusion, or the solar system of the atom. But such an analogy is only a first step; the directive is not simply to “look” at the former. The proffered advice is to start with the equations (and hence, class of mathematical models) that have been used to describe it, in the construction of a theory (hence, a class of mathematical models) of the latter.

<sup>4</sup>Henceforth, “model” will be used in the sense of the scientist (model of a phenomenon, or of the atom, the solar system, the cosmos, Appalachian English, etc.) unless otherwise qualified. Models are structures, usually mathematical structures.

<sup>5</sup>It is also to be contrasted with the sort of psychologism that intrapolates *mechanism of interpretation* between sensation and perceptual judgement. Despite superficial similarities, there is a great difference between the critique of foundationalist notions of the Given in e.g. Reichenbach’s *Experience and Prediction* and C. I. Lewis’ *Mind and the World Order*.

<sup>6</sup>Important for some derivative uses is the idea of *conditional classification*, that is, classification *on the supposition of* (for example, of theories believed by the utterer). This may harbor some logical complexity (think of conditional obligation and conditional probability).

<sup>7</sup>Foley (1993, Chapter 1) advocates an essentially similar account, though in much greater detail, for rationality in general. His discussion provides a good general framework, it seems to me, for Giere’s proposal.

<sup>8</sup>Despite the phrasing, I think this point does not rest on a sharp fact/value distinction. There is no reason to deny that many of the predicates we use resist being classified wholly on either side of that divide.

<sup>9</sup>This analysis of evaluative judgement may however fit certain kinds of moral judgement, as in “The Roman practice of slavery was immoral, regardless of what they themselves believed about it and regardless of its function in their society.”

## References

- Ayer, A.J. (1946), *Language, Truth and Logic*. First published 1935; 2nd ed. New York: Dover.
- Beth, E.W. (1963), “Carnap’s views on the advantages of constructed systems over natural languages in the philosophy of science”, pp. 469-502 in Schilpp (1963).
- Carnap, R. (1928), *Der Logische Aufbau der Welt*. English tr. R. A. George, *The Logical Structure of the World*. Berkeley: University of California Press, 1967.
- (1934), *Logische Syntax der Sprache*. English tr. A. Smeaton, *The Logical Syntax of Language*. London: Kegan Paul, 1937.



- Feyerabend, P. (1958), "An attempt at a realistic interpretation of experience", *Proceedings of the Aristotelian Society*; reprinted in Feyerabend (1981).
- (1981), *Realism, Rationalism and Scientific Method*. (Philosophical Papers, vol. 1). Cambridge: Cambridge University Press.
- Foley, R. (1993), *Working Without a Net: A Study of Egocentric Epistemology*. New York: Oxford University Press.
- Giere, R. (1985), "Philosophy of Science Naturalized", *Philosophy of Science* 52: 331-356.
- (1988), *Explaining Science: a Cognitive Approach*. Chicago: University of Chicago Press.
- Laudan, L. (1990), *Science and Relativism*. Chicago: University of Chicago Press.
- Laudan, L. and Leplin, J. (1991), "Empirical equivalence and underdetermination". *Journal of Philosophy* 88: 449-472.
- Putnam, H. (1982), "Why reason can't be naturalized". *Synthese* 52: 3-23; reprinted in Putnam (1983).
- (1983), *Realism and Reason, Philosophical Papers*, vol. 3. Cambridge: Cambridge University Press.
- Reichenbach, H. (1938), *Experience and Prediction*. Chicago: University of Chicago Press.
- Schilpp, P.A. (ed.) (1963), *The Philosophy of Rudolf Carnap*. LaSalle, IL: Open Court.
- Sellars, W. (1965), "Scientific realism or irenic instrumentalism: a critique of Nagel and Feyerabend on theoretical explanation". *Boston Studies in the Philosophy of Science*, vol.2. Ed. R. Cohen and M. Wartofsky. New York: Humanities Press, pp. 171-204.
- van Fraassen, B. (1980), *The Scientific Image*. Oxford: Oxford University Press.
- (1986), "The World We Speak Of, and the Language We Live In", *Philosophy and Culture: Proc. of the XVII-th World Congress of Philosophy (Montreal, 1983)*. Montreal: Editions du Beffroi.