

PART IV

GENERAL PHILOSOPHY OF SCIENCE (A)

Instrumental Evaluation in Scientific Knowledge

F. John Clendinnen

University of Melbourne

Unlike some recent authors, Hilary Putnam recognizes that we can not avoid inquiring about the normative force of the principles that guide scientific reasoning. His answer is in terms of values. In presenting his case for "Internal Realism", he argues that values are presupposed in statements of fact (1981, pp. 128-134). The central thesis in his argument is that truth is not a correspondence with an "unconceptualized reality" and that "the claim that science seeks to discover the truth can mean no more than that science seeks to construct a world picture which, in the ideal limit, satisfies certain criteria of rational acceptability" (p. 130). We adopt these criteria because having a theory which conforms to them is valuable to us; it is part of human flourishing (pp. 133-134).

We must, however, distinguish two kinds of evaluation; as well as attributing value to goals, we may evaluate means with respect to ends, and it is not immediately apparent that the latter reduces to the former. If it does not, and if science has an instrumental role which is crucial, then it may be that there are evaluations which are implicit in all scientific method but which are quite distinct from the evaluation of goals.

If Putnam is right, there are certain features that a theoretical system can possess and which we value and call virtues. If we were to value different features of systems then we would accept as true statements other than those we do. The only thing that distinguishes the statements we accept from those that we might have accepted is that the former belong to a system which has valuable features. Putnam's account precludes a plausible interpretation. He can not allow that these features are valuable only because they jointly make it more likely that the system has some single, desired property; for this would mean that the former properties were only derivatively valuable. Also it is clear enough that we primarily want the system of beliefs that we accept to be true. So for the values of our method of appraising statements to be only derivative they would surely be suitable means to attaining truth; and this would cut right across Putnam's argument. He insists that our wanting true beliefs is constituted by our wanting a system of beliefs with features which

we value immediately. It is hard to see on this account how these features could be anything other than ultimate values. The thesis that emerges is that there are multiple, independently valuable criteria which constitute truth.

1. Empirical Underdetermination

Putnam's argument builds on the widely held view that there are a number of distinct criteria employed in the appraisal of scientific theory. Empirical adequacy is generally accepted as one, but only one, of these; on its own it is seen as underdetermining theory. Putnam does not accept the simple-minded interpretation of this criterion: namely that there are some statements which directly describe what is given in experience and that theory must be compatible with these. However he allows instrumental efficacy as one of the criteria of rational acceptability, and elsewhere makes it quite clear that there is empirical input to our beliefs (pp. 134&54). The nature of this input is not spelt out in detail, but he clearly allows that what is given in experience determines the success or failure of our goal-directed acts. Thus Putnam's criterion of "instrumental efficacy" can be roughly equated with the "empirical adequacy" of other authors, for instance Bas van Fraassen (1980, Chs. 2-3).

The issue we must consider concerns the status of the criterion of instrumental efficacy (or empirical adequacy). Is it on a par with one or more other values, or does it occupy a more basic status? If the latter, then is it possible to hold that the other criteria are independent and non-derivative values? As well as "being instrumentally efficacious" there are three other features of a theory which Putnam holds to be valued: namely being "coherent, comprehensive, and functionally simple" (p. 134). There can be no doubt that considerations of simplicity and coherence (whether or not they are subsumable under a single principle) are employed in theory appraisal and acceptance. What is not so certain is that in doing this we assign value to these criteria which is separate from and independent of our valuing of empirical adequacy.

If, in a certain kind of appraisal, there are a number of distinct criteria, we must consider the possibility of items which meet some of these but not others. Such cases are in no way problematic when we accept the independence of the criteria. For instance a car may well have good acceleration but poor fuel economy. When, on the other hand, we do find a conflict problematic the suggestion must be that the different criteria are separate indicators of a single property. This certainly seems to be the case with truth; conflicts of criteria are seen as problems that need to be resolved. Putnam may argue that this is due to an ill founded, common prejudice; but if the way we take our concepts counts for anything it constitutes some sort of *prima facie* case that "truth" is a more unified concept than he suggests.

A preliminary reconnoitre also indicates a *prima facie* case that empirical adequacy dominates the other criteria. We certainly seek theories which are as simple and coherent as possible, but we must be

willing to accept very complex theories when necessary. If, however, beliefs lead to experiential expectations which are falsified, something has to give. We can not say that the set of statements is weak on empirical adequacy, although rated true overall because it does very well on the other criteria. We may be uncertain as to which, but one statement must go from a set that gives a false prediction. And while we may deny that any statement directly expresses what is given in experience, we must allow empirical falsification of sets of statements. To do otherwise would be to deny any experiential input to our system of belief. There seems to be grounds for taking empirical adequacy as a necessary, if not sufficient criterion of truth. Are there counter-examples to this suggestion? The closest approximation would be those occasional cases of a theory which we think is closer to the truth than another which works better in practical application. However it is significant that these are comparative cases. We do not say that the unsuccessful theory is, on balance, true. We may believe, for various reasons, that it will be by modifying this theory, rather than the other, that a true one will be achieved, but while the theory is in a form in which it is in conflict with experience it is not true.

Someone defending a position like Putnam's may well insist that our notion of truth as an all-or-nothing attribute is a simplification. Even accepting this, it is possible to argue for the primacy of empirical adequacy amongst the criteria for truth. In so far as this characteristic is defective, to that extent the statement, or set of statements, lack truth. Criteria like simplicity and coherence do not work like that. We want to maximize them as far as possible, but a theory by being complex, or a mere conjunction of distinct theories, does not thereby lack truth. The former criterion is still necessary in a way that the others are not; the latter are merely desiderata which we must seek to maximize.

That one criterion is required in a way that others are not is compatible with each of the set of criteria being independent of the others. So let us explore the possibility that empirical adequacy is a necessary criterion of truth, that this criterion underdetermines the selection of statements and that we accept as true only those statements which meet not only this criterion but also belong to a theory which maximizes the desiderata of simplicity and coherence. Now it would be possible to hold this thesis in the rather attenuated sense that there is a merely logical possibility of the instrumental criterion failing to determine, and that if this were to occur then it would be proper to use the other two criteria. This would not imply the existence of actual cases in which we needed to use the latter criteria. However historical studies of methodology are full of evidence that these criteria are widely used.

If simplicity and coherence are actually needed in selecting between theories, and also these criteria are independent of empirical adequacy, then it follows that there are plenty of examples of competing theories which are equal in their empirical adequacy. What the historical record shows does not actually support this. There certainly are occasions when competing theories are equal in their capacity to account for known facts, but as time goes on one or the

other ends up as superior in this respect. At the time that it is impossible to decide between two theories it is, typically, possible to recognize their different predictive content. If two theories each led us to expect exactly the same outcome of every situation we could imagine, then we would start to doubt whether they were genuinely logically distinct. (There can be cases where the use of basically different mathematical devices can produce formulations which look very different but express logically the same theory.)

The possibility of there being genuinely distinct theories which continue to account equally well for all empirical evidence is a serious issue for scientific realism. For the purposes of this paper it need not be resolved; it is sufficient for us that in the typical case where simplicity and coherence are employed, decisions are made between theories which differ in their predictive content; so as time passes the theories will differ in their empirical adequacy. The criteria are used in the expectation that the theory so selected will be maximally efficacious in experimental predicting and in guiding practical action. If the only goal of science should be empirical adequacy, then these criteria, which some have called 'pragmatic', would still need to be used.

What emerges quite clearly is that empirical adequacy (or instrumental efficacy) is involved in theory choice at two distinct points. Past efficacy can be used as a criterion of choice; however at least one of our goals is that the theory continue to be efficacious; but this goal can not be used as a *criterion* in theory choice. (Whether this is the only goal may be an issue between van Fraassen's Constructive Empiricism and Scientific Realism, but does not bear on the argument here.)

2. Practical Justification

Putnam accounts for the normative status of the intellectual procedures used in science by holding that the criteria we use are ultimate, that is unjustified, values. This approach certainly involves the problems associated with an objectivist theory of value, but it promises the great benefit of avoiding the circularity that seems to bedevil pragmatic justifications. If it is held that a certain method is the rational means to achieving a certain end, then we are obliged to ask what relationship makes this so, and also how we can know that it obtains. Knowing that M will achieve E is the standard way of justifying the use of M with respect to E as goal. But if what we seek to justify is the only method by which we can hope to establish what will achieve what, then we are confronted by a vicious circle. Putnam's approach to rationality avoids this circle, for he holds that we use the criteria we do because we accept them as valuable in themselves. What I have been arguing is that this approach will not do because we use such criteria as simplicity and coherence as the proper guides in choosing theories that we wish to be at least efficacious in empirical predicting. We have no choice but to face the problems of a practical justification of our method of belief-appraisal, for we want our beliefs to fulfil a practical role.

Indeed we can not even avoid the question of why it is reasonable to

choose a theory which has worked up to the time of choice. Once it is allowed that the property we want in a theory is not fully present at the time we choose it and that what we want is a theory that will work in the future, we are confronted by the old-fashioned problem of induction. However Nelson Goodman has shown us new dimensions of this problem (1965, Ch. 3); dimensions which can be used to underline the instrumental role of criteria like simplicity.

Requiring that the laws we use to predict be such that their previous use would have given only true predictions may seem an obvious rational constraint; but unfortunately it places no limitations at all on the predictions we can make. (If you would predict that the next emerald viewed will be blue, or any other colour X, employ the law -- which will be compatible with all we now know directly -- "All emeralds are green up to T and thereafter X" where T is the instant at which the prediction is made. By his ingenious definition of predicates like "grue", Goodman showed how such laws could be expressed in as syntactically simple a form as the more usual "Emeralds are green".) So if induction is a definite way of actually predicting, its specification must consist of more than the above requirement. Goodman argued, in effect, against the possibility of using simplicity as an additional requirement; but he did so by assuming that any acceptable simplicity requirement must be syntactical; and then showed that there was a syntactical symmetry between defining "grue" in terms of "green" (and similar predicates) and vice versa. But consider the simplicity of the way a predicate may be given meaning in terms of experience; there can be no doubt that "green" is simpler in this respect than "grue". Perhaps simplicity in this sense can be used to solve Goodman's "New Problem of Induction". In any case what is quite certain is that some criterion other than predictive efficacy amongst the facts to hand must be used to even have a univocal method of predicting. So there is no question that we can not pursue the goal of predictive adequacy by doing no more than using predictive adequacy amongst the facts to hand as a criterion in selecting predicting rules. We have to consider why simplicity, or some alternative criterion, ought to be used if we would predict correctly. (Goodman offered a relativistic answer which will be unsatisfactory to those who take method to be normative.)

Can the circularity of a practical justification be circumvented? Hans Reichenbach, in his work on the justification of induction, considered this question with some care (1935). He concluded that the justifying argument must be deductive and its premises analytic. Proceeding in this way certainly avoids the threatened circularity, for the principle being justified is not presupposed in the justification. The problem is to find a relationship which can be established by such an argument and which will also constitute the rationality of that means with respect to that end. We can clearly not hope to show the sufficiency of the means for attaining the goal. Reichenbach pointed to the possibility of some more complex relationship (p. 474). The justifying relationship he hoped to use was: that M being a sufficient condition for E was a necessary condition for E being attainable. That establishing such a relationship would be a justification of M with respect to E has not, to my knowledge, been challenged, but the possibility of justifying

induction in this way certainly has (although the issue is too complex to be summarized here). Nevertheless the attempt is interesting, pointing as it does to a possible way of avoiding circularity.

Wilfrid Sellars has proposed a justification via deductive argument, not only of induction (1964), but also of the principles that govern our acceptance of what might be called basic empirical statements. In a recent paper he holds that "achieving a certain end or goal can be (deductively) shown to require a certain integrated system of means" (1979, §68). The goal he characterizes schematically as "being in a general position, so far as in us lies, to act". He stresses here two points that are crucial to the whole question of rationality. Firstly, that in adopting beliefs we are concerned that they be instrumentally efficacious. Secondly, any argument that hopes to show what we ought to do in pursuit of this goal can not be ampliative. To these core points he adds that we can not expect, as a justification, the deduction of any simple relationship between means and end. Rather we must consider our whole complex system of thought and explore its interrelations and the possible alternatives open at each point. Only in this way will it be possible to recognize why the only rational option open to us is to proceed according to certain general principles: like induction, methodological simplicity and attributing a substantial initial probability to our ostensible perceptual beliefs.

Sellars' approach opens up new perspectives. Some will disagree with the details of why he says it is reasonable to accept the basic principles, but it must be allowed that he points to a good deal of territory that remains to be explored. The view that no justification is necessary has often been consolidated by a conviction that all possible avenues have been found to be closed. Positive arguments that a justification is needed are supplemented when we see possibilities for a justification that have not been fully exploited.

Sellars does not spell out fully the deductive justifying argument, but it seems clear that a key notion is the following: "To be one who makes epistemic appraisals is to be in this framework [which includes the basic principles]" (§75). The relationship between epistemic appraisal and goal directed action is only indicated, but this presents no problem. What Sellars does not show is why no other framework could serve as a basis for appraisal, and this is the weak point in his account. In spite of citing the goal of being an agent, and presumably a successful agent, he proposes a justification which only brings out what is constitutive of the framework we *actually* accept and passes over the question why we ought to accept that framework in seeking our goal. It could not be literally true that there is no other framework which could possibly be used for appraisal. What is possible is that there is no other framework which could *reasonably* be used. But this needs to be shown. Traditional attempts in this direction have foundered because they tried to show that the principles to be justified have some positive virtue lacked by all alternatives; and this led to circularity. However there is a possibility that needs to be explored. There may be a notion of irrationality which is logically prior to positive standards of what is reasonable, something like "intrinsically self defeating" (cf.

Clendinnen 1982). If this should prove to be so, what we may hope to show is that there is some defect that must be present in all alternative methods which rules them out as reasonable ways of appraising beliefs which are to guide actions which we hope will succeed. That is, to adapt the language of the passage quoted above, we may seek to prove that to be one who makes epistemic appraisals which do not suffer from a defect which renders them irrational is to be in this framework.

References

- Clendinnen, F. John. (1982). "Rational Expectation and Simplicity." In What? Where? When? Why? Edited by Robert McLaughlin. Dordrecht: D. Reidel Publishing Company. Pages 1-25.
- Goodman, Nelson. (1965). Fact, Fiction and Forecast. Indianapolis: Bobbs-Merrill. 2nd ed. (1st ed., 1955).
- Putnam, Hilary. (1981). Reason, Truth and History. Cambridge: Cambridge University Press.
- Reichenbach, Hans. (1935). Wahrscheinlichkeitslehre. Eine Untersuchung über die logischen und mathematischen Grundlagen der Wahrscheinlichkeitsrechnung. Leiden: Sijthoff. (As reprinted as The Theory of Probability. 2nd ed. (trans.) E.H. Hutten and M. Reichenbach. Berkeley, CA: University of California Press, 1949.)
- Sellars, Wilfrid. (1964). "Induction as Vindication." Philosophy of Science 31: 197-231.
- (1979). "More on Givenness and Explanatory Coherence." In Justification and Knowledge. Edited by George S. Pappas. Dordrecht: D. Reidel Publishing Company. Pages 169-182. (This paper is a revised version of the central argument of "Givenness and Explanatory Coherence." The Journal of Philosophy 70 (1973): 612-624.)
- van Fraassen, Bas C. (1980). The Scientific Image. Oxford: Clarendon Press.