Theories, Concepts and Rationality in an Evolutionary Account of Science

NEIL TENNANT

Department of Philosophy,
Faculty of Arts, Australian National University,
Canberra, ACT 2601, Australia

Hull's project has two aspects — one empirical, the other conceptual. I take them in turn.

THE EMPIRICAL ASPECT

The empirical one involves finding out exactly how scientific theories and conceptual systems originate and are transmitted. One needs to investigate the psychology of the creative process; the social structure of scientific communities; the economics and politics involved in the funding of research programmes; the dynamics of the publication process; the politics and logistics of experimental testing; and like matters. In this connection Hull has provided interesting data and insights, which I would not wish to challenge. He does not, however, go so far as to draw on the work of the Edinburgh school of sociology of science (Bloor 1976; Barnes 1977; Compare Knorr-Cetina 1981; also Pickering 1984); but that may be because he would not wish to endorse their anti-rationalist, relativist conclusions (for more on which, see below). I think Hull's notions of *conceptual inclusive fitness* and of the *demic structure of science* offer the prospect of fertile application and extensions in this area. His analysis of the balance between cooperation and competition among scientists puts many aspects of the scientific enterprise into illuminating focus — especially the practices of mutual citation, priority disputes, and the ethical norms within the scientific community governing fraudulent, plagiarized or shoddy research. I think one of the most interesting implications of Hull's analysis, which he could have emphasized more, is that if science is best understood (descriptively or normatively) as a matter of conjecture and refutation, then the labour tends to be divided: one research group's conjecture tends to be the subject matter of a *rival* research group's attempts at refutation.

THE CONCEPTUAL ASPECT

The other aspect of Hull's project — the conceptual one — is more problematic. Hull puts forward a very general evolutionary model, designed to accomodate both cultural evolution (or what he repeatedly calls "conceptual change") as well as different kinds of selection processes at the biological level. He wishes to avoid the routine sort of analogising that involves projecting *from* relatively well-understood biological models *to* something roughly similar at the cultural level. Instead. he abstracts to a more general level from which one can later descend to recover the particulars of *all* these different processes at *both* the cultural *and* the biological levels.

Now while I have no objection to this procedure in principle, I do have reservations about the way Hull has carried it out. Firstly, I have some quibbles about his generalized notions of *replicator, interactor*, and *lineage*, as they apply to the better understood process

Biology and Philosophy **3** (1988) 224—231.
© 1988 *by Kluwer Academic Publishers.*

of natural selection in populations of biological organisms. In particular, I do not regard *extinction* or *death* as in any way necessary for evolutionary change. Differential perpetuation of replicators can be secured by differential proliferation of interactors even in a population of immortal individuals. True, death speeds the process up; but the wild type can change even without it (Schilcher and Tennant 1984, p. 7). I also do not understand why, if speciation is saltative, Hull refuses to regard species themselves as lineages. (This refusal may be connected with his apparent omission of theoretical change without conceptual change as a process deserving the label "evolutionary". For more on this, see below.) But I think these quibbles could be resolved without undermining broad agreement with Hull's general analysis of selection processes. I also accept his replies to the usual objections one encounters to the "disanalogy" between biological and cultural selection: objections concerning speed, particulateness, polyploidy, cross-lineage borrowing and the supposedly "Lamarckian" character of cultural evolution. (For other points of disanalogy, which nevertheless do not detract unduly from the value of the overall analogy, see Schilcher and Tennant 1984, pp. 118—19.)

But my second objection to his general analysis, of which biological and cultural evolution are meant to be special cases, concerns the unsatisfactory way in which Hull has applied his general scheme to the cultural case. This *may* be the result of abstracting his whole book into this target article; we shall have to wait and see. But, as the paper stands, it suffers two defects:

(a) it is too vague on the crucial question of *what the scientific replicators are*; and
(b) it seems to me to be wrong about *what the scientific interactors are.*

(A) WHAT ARE THE REPLICATORS?

It is not enough to say that

> the replicators in science are elements of the substantive content of science — beliefs about the goals of science, the proper ways to go about realizing these goals, problems and their possible solutions, modes of representation, accumulated data reports, and so on.

Does the phrase "beliefs about" here have wide scope, covering all the items that follow in this quotation? Or are beliefs, *along with* problems, "ways of going about . . .", modes, and reports all examples of replicators? Later Hull says

> conceptual replication is a matter of ideas giving rise to ideas via physical vehicles, some of which also function as interactors

and here, it seems to me, his use of "ideas" is ambiguous between *concepts* and *beliefs*. But then we are reminded that

> replicators must not only exhibit structure but also pass it on to subsequent replicators

and this invites the construal of "ideas" as *beliefs*; for *concepts* can be unstructured, primitive items in a conceptual scheme.

Yet later, in a very important section entitled "The Type Specimen Method of Reference", Hull makes clear that he sees the selection process in conceptual change as involving *term-tokens* which are grouped into lineages and trees by means of transmission. "Term-tokens are the things that are being differentially perpetuated." Here, it seems, he is

working resolutely at the sub-sentential level, inviting the reading *concepts* for his earlier ambiguous talk of "ideas". So in the course of his paper Hull shifts from regarding replicators as (tokens expressing) *beliefs* to regarding them as (tokens expressing) *concepts*.

The difference is important, and no mere philosopher's quibble. For simple repetition of a term — that is, proliferation of term-tokens of its type — does not alone guarantee the right status, within an evolving conceptual scheme, for the concept that the term stands for. (Let us even put aside here the further problem that one and the same token can come to "code for" very different concepts in turn, by being embedded within an evolving framework, or theory, that contributes to the term's explicit or contextual definition.) Even with stable meaning or reference assignments — insofar as these are possible — mere replication of a term is not enough to underwrite its status as a currently important and integral part of a successfully functioning conceptual scheme. The term "phlogiston", for example, *might* (for all I know) have been used *more often* in contexts designed to discredit the theory of combustion based on that notion, than it was in the heyday of the original theory. This example makes the simple but important point that only in the context of sentences and whole theories can the significance and utility of a concept-term be grasped. It is as the containing theories and sentences change that the changes in the concept expressed by the term can be appreciated.

This is not to criticise Hull for having said anything strictly incorrect about conceptual change; it is only to point out that a more full-blooded application of his general evolutionary schema to the case of conceptual change would have taken the context principle seriously. It would not have left "replicator" or "term", in this context, ambiguous between *concept* or *belief*; rather, it would have anticipated the important interdependence of semantic levels (much like the interdependence of different levels of biological organization, which he emphasizes a great deal) and made more generous provision accordingly. So much for vagueness about what the replicators are.

(B) WHAT ARE THE INTERACTORS?

Hull claims that

> (s)cientists are the ones who notice problems, think up possible solutions, and attempt to test them. They are the primary interactors in scientific change.

I think this is to confuse their intentional activity — which no-one disputes — with the grounds for calling something, *within Hull's own schema*, an interactor. Consider, for a moment, an animal breeder. He or she is the one who notices what trait is needed in a certain kind of hunting animal; who thinks up possible solutions (mating this animal with that one); and attempts to test them (by selectively breeding for a few generations, to see if the desired kind of animal can be produced). These considerations do not suffice to make the animal breeder the *interactor* in connection with the replicators (in this case, the genes) belonging to the animals undergoing selection. The connection has to be more intimate than that. And here, I think, Hull's general schema suffers from a deficiency in not requiring a tighter connection between interactors and replicators. The replicators should, in an appropriate sense, *produce* the interactors. This could be in a *causal* sense (the way genes produce organisms, given the appropriate environment); or it could be in a *constitutive* sense (the way axioms produce a theory by means of logical closure — for more on which, see below). Either way, what is important is that *variation* at the level of replicators should show up as *variation* at the level of interactors. In this way the selective filtration of interactors is able to penetrate through to the level of replicators so as to affect their distribution and produce evolutionary effects.

In my view, scientists are just too gross as entities to link up as interactors to the concepts and beliefs serving as replicators. For one thing, scientists *could* — even if, as a matter of historical and psychological fact, they do not — undergo a succession of theoretical allegiances within a single lifetime: taking up one theory, then casting it off, taking up another (possibly radically incommensurable) one, then casting if off, and so on. Different sets of replicators, but same interactor? — surely not. It would be too much like imagining that radically different genotypes could be "tried out" in "one and the same" individual organism, or phenotype, serving as the interactor in Hull's scheme. And not only are scientists too gross in this sense; we see from the example that they are also not intimately enough connected, by either causal or constitutive links, to the replicators for which they are supposed to be the interactors. Finally, it is *not even* (to invoke Hull's own criterion)

> the differential (extinction and) proliferation of scientists that causes the differential perpetuation of the replicators that "produced" them . . . [My parentheses and scare quotes — NT]

Death rates and family size of scientists seem to me to have little or nothing to do with whether a theory or conceptual system get propagated. It seems to me, rather, that if anything to do with scientists themselves is to count, in Hull's schema, as an interactor (for the concepts and beliefs identified as replicators), then it should be *ideological time-slices* of living scientists; or *momentary intellectual profiles*; or *instantaneous memotypes* — call them what you will. These can be suitably construed, given a materialist theory of mind, as *in the physical world*: they will exist as patterns of neuronal linkages and states of excitation, say (Schilcher and Tennant 1984, p. 117 and p. 124). Scientists themselves — the population of actual living bodies — are then just part of the *selective environment* with which *these* interactors, now properly identified as such, interact. No doubt it is a very important part — so important that Hull was mistakenly tempted to regard individual scientists as the interactors.

A REVISED APPLICATION OF HULL'S SCHEMA

I want now, on Hull's behalf, as it were, to suggest what I take to be the most fruitful way to apply his own schema. Then, finally, I shall enter some further *philosophical* reflections on the overall aims and achievements of this branch of evolutionary epistemology.

The suggestion is already implicit in what has been said above. For the sake of simplicity, let us suppose that science is done in a regimented idiom such as a first order language with the usual logical operators: connectives, and quantifiers with variables. The language will have extra-logical vocabulary in the form of names, function signs, and predicates symbols. I know this is an over-simplification of the actual linguistic resources of scientific discourse. It does not allow for the modalities of physical possibility and necessity; for counterfactual locutions, which are so important for a proper understanding of laws of nature as opposed to accidentally true generalizations; for tensed statements (apart from those that can be handled using quantification over times); and many other aspects of logical grammar that could be essential to a completed science and an adequate semantics for its language. But the idealization is not vicious; for all these added complexities could be brought into the picture later, once my general point has been made about the complications, *already present* in the simplified language, that must be dealt with *directly and simultaneously* if Hull's general analysis of conceptual change in science is to have satisfying consequences.

The linguistic simplification could indeed be carried to the point where (by well-known techniques) names and function signs were replaced by predicate symbols. Three levels would then present themselves for consideration:

(1) the level of *predicates* (which express *concepts*);
(2) the level of *sentences* (which express *beliefs*); and
(3) the level of *theories* (that is, *systems of belief*)

There is an analogy here (though I would not wish to make anything of it) with the three levels of genes, genomes and genepools. There may also be an analogy with the three levels of genes, polygenes (for particular traits), and whole genomes — but my task, like Hull's, is not to try to understand the cultural by means of (inevitably deficient) analogy with the biological, but to get the cultural as a specific instance of his general model of evolutionary processes.

Predicates have to feature in sentences in a grammatical way. They contribute their meanings to sentences in a way that is constrained by their mode of syntactic containment; and they derive their own meanings further from the way they feature across many different sentences, applied to give a roughly true description of the world. Sentences in turn go to make up whole theories. We can understand a theory as identified with some basic set of axioms, or as the idealized infinite set of sentences that follow logically from that basis. For definiteness I shall take "theory" on the latter construal.

Philosophers of language will be familiar with the way the theory of meaning has progressively enlarged its focus in order to trace determinants of meaning. First there was Frege's (1884) context principle, to the effect that one should enquire after the meaning of a term only within the context of the whole sentence in which it occurs. Then there was Quine's (1951) emphasis on how our theories stand up to the test of experience only as wholes, making the unit of empirical significance the whole theory rather than any individual sentences within it. But this more radical holistic tendency can be tempered by emphasizing two important intuitions:

(i) it is at the level of *sentences* that the language learner breaks into a language; and
(ii) sentences are made up of repeatable constituents that contribute their meanings in a more or less uniform way to the meanings of sentences that contain them.

These intuitions need not be abandoned even if one accepts further that quite large *sets* of sentences (that is, theories) can help to provide contextual definitions for predicates within their member sentences. (For a more extensive accomodation of molecularity with holism, see Tennant 1987a.)

Now how does this linguistic clarification help with the problem of how to understand conceptual change? I shall begin by observing that it is not only *conceptual* change that one wishes to characterize; there is also *theoretical* change. They are *not* one and the same thing. For one can have a stable set of concepts (that is, a stable meaning-assignment to the predicates of one's language) through a succession of theoretical changes. This would be the case (let us call it [A]) if, for example, there were enough agreement on central principles of a theory to allow different peripheral conjectures to be formulated and tested. We would be in a position to understand the significance of the conjectures in advance only by virture of being able to combine the stable, pre-established meanings of constituent terms to get the meanings of the sentences (put forward as the conjectures) that involve them. This process of conjecture, refutation, revision, renewed conjecture, refutation, revision, . . . constitutes *evolutionary theoretical change* without yet involving one in any sort of *conceptual* change.

In the other direction, there just *might possibly* be conceptual change without any genuine theoretical change, at least from the vantage point of one particular way of judging "sameness of theory". If one eschewed my simplistic way of taking a theory to a be a set of sentences, and chose rather to identify it (in instrumentalist spirit) with just its set of *observable consequences* — this more might be congenial to a "constructive empiricist" like

van Fraassen — then it might be the case that two different sets of sentences, involving very different theoretical terms and principles, had the same set of observable consequences, hence counted as the same theory; yet their choice of concepts, reflected in the different ways their respective predicates were constrained to feature in the implicit definitions they respectively provided, would arguably be very different (Quine 1975). I know that this is a far-fetched possibility, and that it uses a non-standard notion of identity of theory, but it helps at least to underscore my point that we should not be over-hasty in identifying conceptual change with theoretical change.

Yet both kinds of change importantly can and do occur, and both kinds can claim to be evolutionary. The problem is how to accomodate them both as special instances of Hull's schema.

I would suggest that we concentrate our main efforts on cases of type [A]. That is, we should try to understand the dynamics of theory change within one and the same conceptual regime, as it were. This would be analogous to understanding microevolution in the biological case. The more dramatic kinds of change — the Kuhnian (1970) "revolutions" in science — which may involve a dismantling and replacement of a whole conceptual sheme, are then extreme cases of this evolutionary process. They would correspond to genetic saltations, Goldschmitt "hopeful monsters", or extremely penetrating "behavioural spearheads" in the biological case (Popper 1972, especially 'Of clouds and clocks'). Just as the functional significance of anatomical features or behaviours might change dramatically, and therefore the genetic "information" of the genes coding for them be interpreted in a different light, so too in the scientific case the old terms, retained because of the inertia and conservatism of linguistic usage, would take on radically re-defined meanings in the post-revolutionary theoretical context in which they were newly embedded. But the important point is that even in the context of revolutionary scientific change, it is the level of *theory* is the most significant, as the primary level at which upheaval occurs, before the semantic shock waves reverberate down to the level of predicates and associated *concepts*.

Although I have said that I do not wish to make much of the analogy with the biological case, I cannot refrain from pointing out one aspect of my proposed application of Hull's schema that makes it nicely appropriate in the context of talk about genetic "information" in the biological case. Just as genotypes express their information in phenotypes, so too do the deduced consequences of theories (construed as sets of axioms) express ("develop") the information contained in the axioms themselves. It is only by "unpacking" theoretical content via testable, observable consequences, that one reaches that part of the theory — what philosophers of science call its "empirical meaning" — that can be regarded as an *interactor* with Nature (the obvious environment to which theories have to be adapted!). The negative selective effect of a disagreement between deduced prediction and observational outcome is, by *modus tollens*, to weed out and replace some basic axiom or axioms with which the deduction of the prediction began. (The so-called Quine-Duhem Problem is that of how to decide exactly which ones to weed out.) In this way, via modifications of the theory in its axiomatic basis, we get *differential replication*. And the *replicators* here are (tokens of) axioms (or of candidate axioms) of the theory. Moreover, the source of *variation* among the potential replicators could well be random, pending a better understanding of the psychology of the creative process. At least one major thinker in this area, Donald Campbell (1960), regards the process of theoretical innovation in this way as fundamentally Darwinian — that is, as random in its source of variation. Thus yet another way of drawing an analogy suggests itself: *axioms* are like *genes* (or *alleles*); *deduced consequences* are like *phenotypic traits*; and *whole theories* are like *individual organisms*. What is interesting in this way of seeing the matter is that on the left one is dealing only with items at the level of *sentences* (expressing *beliefs*); concepts have yet to get into the picture as any point of focus for one interested in scientific change. I am not, however, denying that important changes, of an evolutionary kind, do indeed take place at the level

of concepts. Indeed, I have tried to sketch how this would be the result of change at the level of theories (sets of sentences) that involve predicates expressing concepts.

The kind of theoretical change that I have likened to microevolution is the subject of intense and fruitful research among mathematical logicians at present (Alchourron *et al.* 1985; Alchourron and Makinson 1982), and Hull would be unwise to overlook it as a potential (and, I think, the most important) instantiator of his proposed general scheme. The research is being conducted, for a start, on first order languages of the kind I limited myself to above; and yet all the complexities of the basic problem are to be found already. The logic and dynamics of theory change even in this regimented language need to be better understood if we are to have efficient (and non-catastrophic) expert systems that can up-date their databases, resolve contradictions between new and old data, and "intelligently" retain or reject more general principles, with which they have been equipped, in the light of the new data. It is not beyond the bounds of imagination that expert systems will be developed that will "randomize" the variant hypotheses with which they might respond to new data that conflict with old; and will apply "selective filters" to weed out the worst ones with obvious tests in salient respects, before tentatively putting forward one of the survivors as a new "conjecture". If the current literature on so-called genetic algorithms (Holland 1975) is anything to go by, the results that may be achieved by such fundamentally Darwinian "thinking machines" could be as impressive as those that we get ourselves, with all our so-called "ingenuity" and "insight".

This prospect immediately raises the questions foreshadowed above about the aims and achievements of this branch of epistemology. Hull's contribution is towards an understanding of the evolution of certain cultural products — concepts and beliefs — rather than the evolution of those biologically based capacities that make us cultural creatures. The distinction is by now familiar to practitioners of evolutionary epistemology. It has been drawn by Vollmer (1983), Schilcher and Tennant (1984, especially Chapter 3), and Bradie (1986). The latter introduces a useful pair of acronyms, which I shall use here. He talks of the distinction between a theory of the evolution of mechanisms (EEM) and a theory of the evolution of theories (EET). A writer like Lorenz (1973) is concerned with EEM; writers like Popper (1972), Toulmin (1972) and now Hull, are concerned with EET.

The main questions posed by philosophers to the new practitioners of EET are these:

- how does EET help us gain an understanding of the *normative* force of our methodological ideas?
- does it tell us how we *ought* to set up our scientific theories on the basis of the evidence?
- does it account for *necessities* in our thought? (Stroud 1981, and in criticism Tennant 1983)
- does it help to *justify* our canons of reasoning, either deductive or inductive? (Tennant 1987b)
- does it help us to understand better the notion of *scientific progress*, or *approximation to the truth*? (Oddie 1986)
- does it provide *criteria of choice* between competing theories?
- does it help resolve, or does it serve to reinforce, claims about the *incommensurability* of competing theories or theoretical frameworks?

All these questions can be summed up into one brief one:

- does EET help illuminate the notion of *scientific rationality*?

Now unfortunately I do not find much on offer in Hull's paper by way of an answer to any of these questions. Indeed, he seems to me to make a crucial sidestep of the main issue when he says

> . . . the sort of rational selection of beliefs in which people engage (when all else fails) and all the semiconscious and unconscious selective retention that characterizes how human beings acquire their beliefs are also fundamentally the same sort of phenomenon.

Whether or not we are here to interpret "semiconscious and unconscious selective retention" as rational, it is clear from the context that Hull is not offering his account of evolutionary change as revealing, in any constitutive sense, what it is to be rational. Rather, he is presupposing an understanding of what it is to be rational in simply stating that such-and-such a rational process, along with such-and-such other processes (be they rational or not) are both examples of the sort of general phenomenon he is out to characterize with his schema. My question is: will we have made any progress with the important philosophical questions if we rest content with tracing term-token trees of transmission from an external vantage point, and interest ourselves neither in

(i) the detailed *grounds* (in individual intellectual biographies) for their differential production of term-tokens, nor in
(ii) the questions *whether* and *how* (if his indeed be the correct *causal* picture of conceptual and theoretical change) those changes conform to the dictates of rationality?

These are deep and difficult questions, and one would of course be expecting too much of Hull if one were to demand that he treat them at any length in what is already a long paper full of interesting ideas and discussions of what in the context were more pressing objections. But I cannot help feeling that the ultimate value of the causal stories told after his fashion will be appreciated only after more thought has been given to the possible *emergence* of the normative and the rational from the descriptive and the causal. I myself believe that an evolutionary approach is the key to understanding how "is" yielded "ought". I cannot share the simplistic relativism of the Edinburgh school who see the social character of the selection process as undermining the supposed rationality of its constituent steps and its products. (Nor, do I believe, would Hull.) But then we have to address ourselves to the questions posed above. The main idea I have in an attempt to answer them is this: When language and our conceptual capacities themselves emerged, the evolutionary process was such as to confer on the emerging structures certain meanings, with respect to which their subsequent exercise had to be *faithful* in order to achieve the very ends for which the selection of those structures was taking place. In this way, for example, one can account for the incorporation of logical operators into a growing language, and justify the logic that they obey. This is not the place to develop the details any further; I can only refer the reader to other places where I attempt to do that (Tennant forthcoming). But I am constrained to place on record, by way of comment on Hull's fascinating and suggestive paper, that the connection between his causal, descriptive story, and the story about norms and rationality that philosophers wish to tell (or be told), is one that he will have to investigate in more detail if he wishes to interact successfully with the wider epistemological environment and see tokens of his new schema replicate to good effect in the brains of other epistemologists who are both co-operating and competing with him.