

Are Scientists Agents in Scientific Change?

CECILIA HEYES

*Department of Experimental Psychology
University of Cambridge
Cambridge CB2 3EB, England*

Many proponents of an evolutionary or-selectionist epistemology are beginning to think that they must shortly “put up, or shut up”, but Hull is unusual because he has “put up”, attempted to cash one of evolutionary epistemology’s promissory notes by suggesting how it might apply to conceptual change in science. Responding to this attempt is perhaps more difficult for one who is sympathetic toward selectionist epistemology than for one of its opponents. While an opponent who finds fault with Hull’s account can conclude that they were right all along, that evolutionary epistemology is bankrupt, a sympathiser or protagonist who is critical might justly be expected to provide an alternative to Hull’s proposals. I am a sympathiser, I find several aspects of Hull’s account unsatisfactory, and I cannot offer a comprehensive alternative. However, I offer the following response to the target article because what I criticise are inconsistencies in Hull’s approach, elements that I see as endangering by contradiction his genuine insights. As a consequence, I consider the skeleton of an alternative account to be embodied in the original.

In his paper, Hull seems to set himself two tasks: First, in the introduction and section on ‘Conceptual inclusive fitness’, he seeks to explain the observed behaviour of scientists — their priority disputes, citation patterns etc. — in terms of the motivations of individual scientists, and their small group social organisation. Second, in the remainder of the piece, Hull attempts to identify the components of scientific change necessary for it to qualify as a selection process, and to recommend a methodology for historians of science based on this evolutionary analysis. I regard these problems as distinct both because a correct solution to the first may have few implications for an adequate resolution of the second, and because Hull uses evolutionary theory in a different way to address each problem. In order to explain the behaviour of scientists Hull reasons by analogy from biological or gene-based evolution, while his attempt to identify replicators and interactors in conceptual evolution represents, in his own words, “a general analysis of selection processes” (p. 134); the use of concepts abstracted, rather than directly transferred, from the study of gene-based evolution.

I will discuss each of these two components of Hull’s contribution in turn. In commenting on the first, I will question whether Hull could really have the data necessary to give a satisfactory account of the behaviour of scientists in terms of their mental states and social structure. In connection with the second, I will query Hull’s claim that “individual scientists are the agents in scientific change” (p. 140), on the grounds of its inconsistency with both his premises (components of his general analysis of selection processes) and his professed aim (the specification of a *mechanism* for conceptual evolution), and suggest an alternative view of the identity of conceptual replicators and interactors.

METHODS OF STUDYING SCIENTISTS’ BEHAVIOUR

Hull makes the structure of his first task, as he sees it, fairly clear. He claims that science is successful in that scientists, more than other professionals, behave in accordance with their

own norms, and seeks to replace the traditional view, which attributes this success to disinterested striving after truth on the part of scientists, with an explanation of the successful behaviour in terms of conceptual inclusive fitness and the demic structure of science. Both of these conceptual resources are drawn directly from the evolutionary theory of gene-based selection processes, and the mechanisms to which they refer can be more succinctly labelled as 'credit' and 'checking'. Hull claims that the successful behaviour of scientists is a product of competition for credit (that scientists strive 'to get their views accepted *as their views* by other scientists' (p. 126 — emphasis added), tempered by the checking of results which inevitably occurs when groups of scientists are working on a problem in parallel.

So much for Hull's agenda and conceptual tools, what about his empirical resources? Of course (and it is a fact that the author himself bewails) we are told almost nothing about them in this paper. However, assuming that Hull is a conscientious, naturalised philosopher of science (perhaps the most conscientious), and using grapevine fragments of information about his work, it is possible to piece together the following. Hull has made an intensive study of zoological systematists, past and present, examining their published works, notebooks and correspondence, and, perhaps, informally observing and interviewing some of them at scientific meetings and in their places of work. Naturally, it is impossible to tell without seeing Hull's data whether it supports his conclusions, but as a psychologist I doubt that it could. In view of the importance that Hull himself assigns to the empirical component of his enquiry, I will attempt to communicate my misgivings by mentioning some of the methodological problems that Hull shares with social psychologists, and examining one of his claims in the shadow of these concerns.

The task that Hull sets himself in the first part of his paper resembles those tackled by many social psychologists in a couple of important respects. He is seeking the causes of people's behaviour in their social organisation and in their mental states or dispositions; and both practical and ethical considerations prevent him from executing carefully designed experiments to discover the precise nature of the causal relationships in question. Many of the methodological problems that arise when a social scientist is unable to intervene in, or manipulate, their subject matter, can be grouped under the headings "sampling" and "measurement", and failure to resolve these problems can threaten, among others, the "external validity" and the "construct validity" of the study's conclusions. (This is the terminology of "quasi-experimentation", and I will draw heavily on the authoritative text on this subject, Cook and Campbell 1979, in what follows.)

Consider the relatively simple case of a social psychologist who is being paid by a commercial company to discover whether aggressive sales personnel are better or worse at selling their product than non-aggressive sales personnel. A sampling problem that might threaten the external validity of the study would arise if the psychologist were allowed to test only those sales people who volunteered to be part of the study. Volunteers might be systematically younger and more enthusiastic about their work than their non-volunteering peers and, as a consequence, any conclusion about the relationship between aggression and sales based on a study of this sample of people, could not be assumed to hold for the whole population of the company's sales employees.

Similarly, Hull's study is likely to have been smitten with a couple of sampling problems. He has tested systematists whose behaviour many be unrepresentative of the population of scientists as a whole both because they do not conduct experiments in any conventional sense, and because they rely on the type-specimen method of coding their nomenclature. Since, as Hull tells us, this method stresses priority, they are likely to have been selectively exposed to a working environment that fosters the desire for credit. Thus, the external validity of Hull's conclusion that *all* scientists who exhibit successful behaviour are motivated by a desire for credit is under threat.

The second sampling problem that is likely to have marked Hull's study has the potential to threaten both the internal and the construct validity of some of his conclusions. Using an historical data base, it is almost inevitable that the people one studies will have

received recognition for their work. This being the case, Hull's claim that scientists' desire for credit is part cause of their high frequency of publication of genuine findings is under threat. Its construct validity is dubious (either the hypothetical cause x or the hypothetical effect y could have been confused, or confounded, with some conceptually similar construct, a or b), because the effect of desire for credit (x) on frequency of publication of genuine findings for which the author received credit (b) has been measured, not the effect of x on y , the frequency of publication of genuine findings. The internal validity of Hull's claim is dubious (the causal arrow may be in reverse) because it is plausible that "ontogenetically" the publication of findings for which a scientist is given credit, has the effect of making that scientist desire credit. Hull acknowledges this latter ambiguity in his paper.

Finally in this vein, a problem that is more directly concerned with measurement threatens the construct validity of the same one of Hull's claims. Just as a social psychologist might confound "aggression" with the conceptually similar construct of "assertiveness" if he or she only measure aggression according to how many times a salesperson corrects his or her client, Hull may confound "desire for credit" with, for example, "desire for use" if he does not attempt to triangulate on the construct by employing a range of measures. I focus on the possibility that Hull has confounded this particular pair of constructs not only because they would be associated with similar behaviour, and are therefore eminently confoundable, but also because the hypothesis that scientists generally strive simply to have their views accepted, rather than accepted *and* accredited to them, has several things to recommend it as part of Hull's scheme. Before concluding this section I will discuss these.

First, if desire for use were substituted for desire for credit in Hull's theory, a more consistent analogy between genetic and conceptual inclusive fitness would be sustained. As Hull puts it, "... organisms behave in ways which result in replicates of their own genes or duplicates of these genes in close kin being transmitted to later generations" (p. 126). They do not, and could not, behave in ways which result in replicates of their own genes being transmitted to later generations *in a form which would identify the genes as having once been possessed by that particular organism*. This being the case, it seems that the first hypothesis thrown up by analogical reasoning is that scientists will behave in ways calculated to get their views accepted (that if they desire anything, it is the *use* of their views), and that this should be abandoned only in the face of very strong evidence indicating that they desire credit, i.e. "behave in ways calculated to get their views accepted *as their views* by other scientists" (p. 126 — emphasis added).

Second, it is Hull's claim that scientists desire credit rather than use that creates the anomalies that he spends much of the first half of his paper attempting to explain. Both the desire for credit and the desire for use views would predict that lying (fabricating data) will be punished more severely than stealing (accrediting the views of another scientist to oneself). However, only the desire for credit view, through its implicit denial that scientists apply everyday moral values in their work, would predict that scientists will (i) object minimally or not at all to the theft of *other* scientists work, and (ii) pay little heed to whether the falsification was deliberate when imposing sanctions for lying. Hull cites the Burt case in support of these predictions, but as far as I can see it has no bearing on the first, and actually constitutes counter-evidence with respect to the second.

Commenting on the case Hull suggests: "When other scientists thought that all he [Burt] had done was appropriate to himself the work done by his assistants, no one was especially excited. After all, that is what assistants are for. But when it began to appear that he had fabricated not only these assistants but also their work, his fellow scientists became more than a little anxious because it brought into doubt all of the work that they had published which was based on his fabricated results" (p. 130). It is correct to regard the scandal as having had two phases, mild criticism and concern followed by a veritable furor, and to identify the transition with the discovery that Burt's "assistants" did not exist, or did not exist in the places or at the times necessary for them to have collected the relevant data (Gillie 1976). However, prior to this discovery no one could have reasonably supposed

that the data had merely been stolen. It was the realisation that Burt's correlation coefficients could not have been calculated correctly from *any* genuine data set that initiated the debate (Kamin 1974).

It is more likely that commentators were relatively unperturbed prior to Gillie's revelations because, although they knew that Burt's data were flawed, they assumed that this was due to the carelessness of an old man, rather than to deliberate deception. Jensen (1974) certainly offered this interpretation, and if one were to doubt the representativeness of Jensen's view on the grounds that, having used Burt's data, he had much to lose through its wholesale rejection, then a further difficulty emerges for Hull's analysis. It seems that Jensen and a number of other psychologists who had used Burt's work on the heritability of IQ were in a position to know the extent of Burt's transgressions long before other scientists (Clarke and Clarke, 1979). If, as Hull suggests, scientific lying is punished principally because of the harm that the false data can inflict on their users, then why hadn't these psychologists ceased to make use of Burt's work, raised the alarm themselves, or at least dropped their allegiance to Burt at the first whiff of scandal?

This issue aside, the Burt case provides ample evidence that scientists care whether data have been falsified intentionally or unintentionally. When the falsity, and therefore potential harmfulness, of Burt's correlations had been established beyond reasonable doubt, investigators who had not been harmed continued laboriously to probe the fraud vs carelessness issue. McAskie (1978), for example, scrutinised Burt's data for telling signs of digital preferences, and Clarke (personal communication) initiated correspondence with a man who claimed to have been tested by one of Burt's assistants, just in case the extent of his deliberate deception had been overestimated.

A desire for use interpretation of conceptual inclusive fitness might also facilitate an account of citation. The desire for credit view makes not only certain citation patterns difficult to explain, but also the mere fact of citation an anomaly. As a cost of the conceptual inclusive fitness view of science, this may be acceptable when the only alternative is the implausible claim that scientists seek knowledge for its own sake. However, the desire for use interpretation of conceptual inclusive fitness allows that scientists are in some sense selfish, that they seek a variety of immortality, without incurring such a heavy burden of anomalies.

Of course, if the data favour a desire for credit interpretation, then these pragmatic considerations must be eschewed, but a massive, systematic and subtly design study would be necessary to make the data legislate in this way. It would be helpful if these studies included samples of scientists that vary not only in their disciplinary affiliations, but also in the extent to which they have both desired and received credit for their work. Furthermore, it would be necessary to employ methods of measuring the extent of an individual's desire for credit which are both independent of that individual's achievement of credit, and capable of distinguishing desire for credit from desire for use. As Hull is well aware, simply asking scientists what they want is not an adequate instrument, since they are likely to respond in accordance with their own implicit or explicit theory of science (see Nisbett and Wilson 1977) for evidence supporting this interpretation of 'introspective' reports), but a carefully designed and standardised questionnaire may be capable of disambiguating these constructs. Of course, questionnaires cannot be administered to dead scientists, but this only serves to emphasise the need for studies of contemporary scientists to augment an analysis based on episodes in the history of science.

INTERACTORS AND AGENCY

In the second part of his paper Hull identifies "elements of the substantive content of science" as conceptual replicators, and scientists both as "vehicles for replication sequences" (p. 140) and as interactors. He defends the claim that scientists function as interactors by

pointing out that “Without scientists, no conceptual replicator could ever be tested, and testing is essential to science”; and asserts, apparently as a consequence of their functioning as interactors, that “individual scientists are the agents in scientific change” (p. 140). I do not doubt that testing is an essential component of science, that the operation of scientists’ cognitive machinery is necessary for that testing, and that scientists, like other people, are in some contexts “agents”. However, I find Hull’s identification of scientists as interactors inconsistent with his general analysis of selection processes, and since it appears to be their cognitive complexity, rather than the content of their mental states *per se* that drives conceptual change, I think that it will prove counterproductive to treat scientists as agents in the context of scientific change.

The identity of Hull’s conceptual replicators and interactors is incompatible with his general analysis of selection processes in two respects: (i) elements of the substantive content of science (conceptual replicators) do not *produce* scientists (conceptual interactors); and (ii) it is not the differential extinction and proliferation of scientists (conceptual interactors) which causes the differential perpetuation of elements of the substantive content of science (conceptual replicators). This is the case whether one regards the extinction of scientists to consist in their death or their withdrawal from scientific activity.

If the identification of scientists as conceptual interactors carries with it the notion that scientists are agents in conceptual change, then I would raise a further objection to his identification: Hull perceives the problems which he confronts in this paper as closely allied to those which Darwin tackled throughout his career, and indeed when, in the first part of the article, Hull attempts to overcome anomalies in the behaviour of scientists, their kinship is apparent. However, in applying his general analysis of selection processes to conceptual change, Hull seems to depart radically from Darwin’s approach. While one of Darwin’s principal achievements was to remove the notion of agency from the explanation of adaptation in organic form, Hull has reinforced the significance attributed to it as an explanation for progress in science. In a manner that is apparently inconsistent with his views on the role of intention in conceptual change (expressed both in the target article and in Hull 1980), Hull has challenged the content of the beliefs and motivations commonly attributed to scientists, but he has not rejected the received view that mental states are critical components of the mechanism of scientific change.

To claim that scientists are agents in scientific change implies that the content of their beliefs about how science should be conducted, and their aspirations concerning the pay offs of scientific activity, are an important determinant with respect to, not only the course, but also the fact of scientific progress. I find this emphasis inappropriate for three reasons: First, as Hull has pointed out, within a fairly broad range, which problems scientists intend to solve, or what they believe to be the proper way to do science, has little influence on whether or what they discover. If this were not the case, then scientific interest in the philosophy of science throughout the last century would have done more harm than is apparent, and instances of serendipity would be less pervasive in the history of science. Second, it would be unfortunate if an evolutionary analysis of scientific change were crucially dependent on our understanding the beliefs and motivations of individual scientists since, as I hope that the first section of this commentary illustrates, the content of these states is very difficult to specify empirically. Finally, while it has long been suspected that difficulties in the empirical identification of mental states are but a symptom of their incompatibility with (other) scientific constructs, recent developments in psychology and naturalistic epistemology have clarified the relationship between the “folk” theory of mind of which they are a part and other ‘scientific’ theories of mind, and made their dispensability in the explanation of behaviour more plausible than it has probably ever been before (e.g. Churchland 1979; Dennett 1987). If one takes these developments at all seriously, then one will avoid treating mental states as explanatory constructs wherever possible, and in applying a general analysis of selection processes to conceptual evolution I think that it is possible.

As an alternative to Hull's identification of conceptual replicators and interactors, the former might be regarded as those elements of the substantive content of science that are encoded in *neural* vehicles, and the latter as the same expressed in speech, text, diagrams, gestures and the like. The advantage of this interpretation is that it conforms more closely to Hull's general analysis by making conceptual replicators (i) capable of copying themselves, (ii) producers of conceptual interactors, and (iii) dependent on the differential extinction and proliferation of conceptual interactors for their differential perpetuation. Of course, conceptual replicators ('ideas' embodied in brains) would not be able to copy themselves or to manufacture conceptual interactors (items of text, speech etc.), and conceptual interactors would not be able to function as such, if scientists' cognitive machinery were not of the nature and complexity that it is. However, if this fact is sufficient to make scientists' agents in scientific change, then host organisms must be the agents of change in viruses.

In asserting that speech and script "retain the structure of initiating ideas" (Hull 1988) and thereby denying that there is an analogue of the genotype/phenotype distinction in conceptual evolution, Hull has effectively raised an objection to the present characterisation of replicators and interactors in conceptual evolution. While I have contributed to a discussion of this issue (Heyes and Plotkin forthcoming), I am not fiercely committed to the idea that conceptual replicators and interactors can be distinguished as memes encoded in brains vs memes expressed in books etc. If this seems too arbitrary then it may be necessary to consider the possibility that the *same* entities function as conceptual replicators and interactors. There is, after all, a precedent for this in gene-based evolution, and it would be a consistent position for Hull to adopt given his denial that there is a conceptual genotype/phenotype distinction. The claims that I would like to have stressed are: (1) While the same cognitive characteristics may lead us to regard both people in general as agent, and scientists in particular as necessary for conceptual evolution, it is unlikely to be useful to regard scientists as functioning as agents in conceptual evolution. (2) Scientists with certain cognitive characteristics may be necessary in order for conceptual interactors (elements of the substantive content of science, encoded in books etc., and perhaps brains) to function as such, but while an interactor is defined as "an *entity* that interacts as a cohesive whole with its environment in such a way that this interaction causes replication to be differential", rather than as a *system* that operates such as to make replication differential, then scientists themselves cannot be cogently identified as interactors.

The Mechanisms of Communal Selection and Serendipitous Discovery

AHARON KANTOROVICH

*The Institute for the History and Philosophy of Science and Ideas
Tel-Aviv University, Israel*

1. EPISTEMOLOGY: THE INDIVIDUAL VS THE COMMUNITY

Science is a knowledge-producing institution which reveals to humanity new domains of reality. We would expect, therefore, a philosophical theory dealing with the social development of science to shed some light on the epistemological significance of science. We