I

PERFECT RATIONALITY: BEYOND GRADIENT-CLIMBING

I.1 INTRODUCTION

The aim of this chapter is largely polemical and critical. I shall argue that in spite of certain superficial analogies between the social and the biological sciences, there are fundamental differences that make it unlikely that either can have much to learn from the other. The difference, essentially, lies in the distinction between the *intentional* explanations used in the social sciences and the *functionalist* explanations that are specific to biology. Donald Davidson[1] has argued that the attempts at psycho-physical reduction are bound to hurt themselves on the stumbling-block of intentionality, and from a very different point of view a similar thesis is presented here with respect to psycho-biological reduction or analogy. The argument cuts both ways, against the use of sociological methods in biology and against the transfer of biological paradigms to the social sciences. Most of the space is devoted to the first problem, because it is less extensively discussed in the literature. I argue that even if natural selection can to some extent simulate intentionality, there are crucial differences in the fine grain of animal and human adaptation. Two such differences are discussed in some detail. In I.2 and I.3 a case is made for the capacity for *global maximization* being a specifically

[1] Davidson (1973). The characteristic feature of Davidson's view is that he is simultaneously materialist, determinist *and* anti-reductionist. The following passage brings out the essence of his view: 'If a certain psychological concept applies to one event and not to another, there must be a difference describable in physical terms. But it does not follow that there is a single physically describable difference that distinguishes any two events that differ in a given psychological respect' (p. 717). That is, to conclude from determinism and materialism to reductionism is to commit the same fallacy as when from the fact that everything has a cause one concludes that there is something which is the cause of everything.

1

human trait not found in natural selection. In I.4 a similar argument is made concerning the unique human capacity for *strategic behaviour*. In I.5 I turn to the second problem, drawing attention to the better-known difficulties of functionalist analysis in sociology. Briefly stated, my contention is here that in societies there is no general mechanism – corresponding to natural selection – that could permit us to infer that the latent functions of a structure typically maintain the structure by feedback.

Discussions of these matters have shown me that my argument lends itself to three misunderstandings. First, I do not deny that interdisciplinary research, pooling the empirical and theoretical resources of biology and of sociology, can be fertile indeed. In human evolution, for example, the social structure is part of the environment that determines whether a given mutation will be beneficial or harmful. At another level the biological components of language, intelligence or mental illness are by now well documented, even if the relative importance of nature and nurture is always a matter of controversy. My argument is directed only at the transfer of whole explanatory paradigms; at the appeal to intentions in biology and to functions in sociology.

Secondly, I do not deny that cases can be found where biological adaptation leads to deviation from local maxima and to attainment of global maxima, nor that functional explanation can sometimes be useful in sociology. I am only arguing that there can be no *general* presumption that global maxima will be attained by natural selection, and that latent functions do not *typically* explain the persistence of the social structures exhibiting these functions. We may invoke here the Kantian notion of a *regulative idea*, as opposed to a dogmatic one.[2] A regulative idea distributes the burden of proof, in the sense of postulating which features can be assumed as a working hypothesis in the absence of specific evidence to the contrary and which features must be shown to be present in each particular case. The statement 'Everything has a

---

[2]   *Kritik der reinen Vernunft*, B 670ff (A 642ff); *Kritik der Urteilskraft*, §70ff. The simplest example is perhaps Kant's reinterpretation of the 'principle of continuity' from a dogmatic assertion to the effect that the set of organic forms can be ordered in a continuous (or dense?) series, to a regulative principle that exhorts us to seek, between any two forms, an intermediate one. The statement 'Everything has a cause' is in Kantian philosophy a constitutive principle to which there can be no exceptions, whereas the idea that 'everything has a function' is a regulative one that does admit of exceptions, as is explained in n. 83 below.

function' should not be taken as a dogmatic statement about organisms,[3] but as a statement to the effect that we may legitimately assume that a given feature has a function – i.e. could not undergo small variations without loss of reproductive capacity for the organism concerned – until the contrary has been demonstrated. The justification for this regulative idea is, of course, the general mechanism of gradient-climbing by natural selection, as explained in I.2. On the other hand there is nothing that could justify the regulative idea that every feature of the organism realizes a *global* maximum, so that all variations – small or large – would imply a loss of reproductive capacity. For intentional adaptation exactly the contrary is true. Here we do have a general mechanism for attaining global maxima, and what needs a separate explanation is rather the failure to achieve this.

Thirdly, I would not want to be read as an anti-reductionist, at least not of the extreme variety that postulates an unbridgeable gap between animal and human adaptation. I believe that the human capacity for global maximizing must ultimately be explained as a result of the locally maximizing gradient-climbing in natural selection. In this sense I am indeed a reductionist. Observe, however, that the object of the reduction is the *capacity* to behave in this manner, and not specific cases where this capacity is exercised. I believe, that is, that rational behaviour must be reduced in two steps: first by subsuming it under the general capacity for rational problem-solving and secondly by explaining that *general* capacity by the workings of natural selection. The argument is similar to the one proposed by Peter Richerson and Robert Boyd[4] concerning the biological foundations of the *capacity* for culture, which in their model is very different from the biological foundations for cultural *behaviour*. By contrast the argument is clearly incompatible with the views proposed by C. D. Darlington[5] or R. D. Masters,[6] who tend to see *specific*

---

3   For a clear statement of this dogmatic (or non-non-Darwinian) view we may take the following: 'A "functionless site" is simply one the function of which has not yet been determined' (Mayr 1970, p. 127). A statement along the lines sketched in the present essay is the following: 'The experimental study of adaptation has unravelled adaptive values in such unobtrusive and inconspicuous details of organismic organization that one should think of a character as having survival value until the contrary has been demonstrated' (Curio 1973, p. 1049).
4   Richerson and Boyd (forthcoming).   5   Darlington (1969).
6   Masters (1975, 1976).

behavioural patterns in biological perspectives. I return to the specific–general distinction in I.3.

As a last introductory remark I would like to put forward what I hope may turn out to be a self-destroying prophecy. I do not believe that my criticism, should someone happen to read it, will have any great impact. The attraction of biological analogies on social scientists, in particular, seems to be so great that even the best minds are led astray. In the cabinet of horrors of scientific thought there is room, alas, not only for a Worms or a Lilienfeld, but also for a Durkheim or a Merton.[7] At the present moment I think one can confidently predict that ethology and sociobiology between them will generate a continuous stream of pseudo-scientific papers for some years to come. I do not deny, of course, the occasional utility of biological analogies as a source for new hypotheses, any more than I would deny that some scientists may get their best ideas when reading the Bible or the *Dialectics of Nature*. I do deny, however, that biological analogies should have any privileged status. Ideas should be judged by their descendants, not by their ancestors. I hope that the social scientist who proposes to spend his time on finding *the* social analogue of, say, dominance or flight behaviour in animals would first reflect on the pseudo-debates of the nineteenth century about *the* correct social analogue of the cell: individual or family? There is no reason why any such analogue should exist, and that's all there is to the question.

I.2   THE LOCALLY MAXIMIZING MACHINE

In this section I sketch a highly simplified account of the theory of natural selection. By keeping strictly to first principles I hope to avoid being too patently wrong on specific biological matters outside my competence. As the conclusions themselves will be on the level of first principles, I believe that my simplifications can be justified.

I conceive, then, of the organisms in a population as a machine receiving inputs in the form of mutations. For simplicity we assume asexual reproduction, so that mutations are the only source of genetic novelty. Alternatively we may skirt recombination by

[7]   For accounts of biological models in the social sciences see Stark (1962), Schlanger (1971), Banton (ed.) (1961).

I.2   *The locally maximizing machine*   5

arguing that in the long run only mutations can disturb the
biological equilibrium (in a constant environment). The stream of
inputs is random, in the sense that there is no correlation between
the functional requirements of the organism and the probability
of occurrence of a mutation satisfying these requirements. By
mutagens such as X-rays it is possible to increase the probability
of mutations generally; by chemical mutagens it is possible to
increase the probability of structurally defined subgroups of
mutations, such as the probability of cytosine mutating to thymine
as a result of deamination by nitrous acid; but it is never possible
– and this is the *central dogma* of molecular biology – to modify the
probability of functionally defined subgroups of mutations. An
analogy may help here. Comparing mutations to misprints, we may
increase the probability of misprints by breaking the glasses of the
typesetter, but there is no way of selectively increasing the
probability of misprints occurring in the second edition of a book
that will correct the factual errors of the first edition.

We assume – crucially – that all mutations are small, being
typically amino acid substitutions resulting from the misprint of
a single letter in the genetic code. There are no doubt mechanisms,
such as gene duplication, that can produce macro-mutations, but
in the first place the evolutionary importance of these is far from
being clear and, in the second place, such mutations, while large
compared to amino acid substitutions, are small compared to the
discontinuities that are found in human adaptation.[8] No gene
duplication could produce a change of the order of magnitude of
the switch from the horse-drawn carriage to the 'horseless'
carriage.[9] As the main contention of the present chapter is that
there is a basic difference between the local optimization through

[8]   For the first point, see Frazzetta (1975), pp. 93ff; for the second, *ibid.* pp. 20, 152. For
one possible evolutionary consequence (*not* function!) of gene duplication, see Rigby
et al. (1974), who argue that the duplicated gene can mutate to a functionally inactive
one (one step backwards) and then further mutate to a viable and possibly superior form
(two steps forwards).

[9]   Frazzetta (1975), p. 152. Cp. also Schumpeter (1934), p. 54: 'Add successively as many
mail coaches as you please, you will never get a railway thereby.' For an alternative view
on technological development, stressing the small cumulative changes, see Rosenberg
(1976), pp. 66, 166 and *passim*; David (1975); Nelson and Winter (1974, 1976); cp. also
nn. 27 and 28 below. Even if we accept, however, that many evolutionary changes are
larger than one-step mutations, and that many cases of technical change are
incremental rather than discontinuous, I believe that there remains a genuine
difference that suffices for my case. Another important difference (Frazzetta 1975, p.
20) is that in evolution there is nothing corresponding to 'useful failures' in
engineering. Evolution never learns from past mistakes.

6   I.2   *The locally maximizing machine*

small improvements and the global maximization that permits
steps of any size, the precise definition of 'small' is not really
important.

In the phrase of Jacques Monod, natural selection operates by
chance and necessity.[10] While mutations are random, the selection
process is deterministic, in the sense that the machine at any given
moment has well-defined criteria for accepting or rejecting any
given mutation. (This means that I shall not deal with genetic drift
and with the problem of non-Darwinian evolution.) The mutation
is accepted if the first organism in which it occurs benefits in the
form of higher reproductive capacity. Since the organism will leave
more descendants than other organisms, the new allele will spread
in the population until it is universally present. (This statement
implies that I shall not deal with frequency-dependent selection
and other sources – such as heterosis – of stable polymorphism.
See, however, I.4 for a discussion of polymorphism as related to
mixed strategies.) Among the results of an accepted mutation is
that the criteria for accepting or rejecting new mutations will
change. The organism is now in a different state from what it was
before the mutation, and may be harmed or benefited from
different inputs. To put the matter briefly: the machine says Yes
or No to each input according to criteria that change each time
it says Yes. If the machine ever arrives at a state in which it says
No to each of the (finitely many) possible inputs, we say that it has
reached a local maximum. The population climbs along a
fitness-gradient until it reaches a point from which all further
movement can be downward only; and there it comes to a halt.
For a given initial state, several local maxima may be accessible,
the choice between which depends upon the random order in
which the mutations happen to occur.

A further analysis of this *locally maximizing machine* must take
account of the possibility of environmental change. If the environ-
ment changes, the criteria for saying Yes or No to mutations
will typically also change. A mutation is not beneficial or harmful
in itself, only with respect to a given genetic background (itself the
outcome of previous mutations) and a given environment. With
a changing environment it may very well be the case that even
instantaneous local maxima are never attained, if the organism
cannot keep pace with its surroundings. The notion of an

[10] Monod (1970); see also Schoffeniels (1973).

'environment' is, however, ambiguous. In the first place environ-
mental changes may refer to changes in the geological or
climatic conditions to the extent that they cause evolutionary
change without being themselves affected by it. (The last proviso
is required to exclude endogenous climatic changes, such as the
change in the atmosphere generated by the evolution of plants.)
In the second place some parts of the environment are themselves
evolving organisms or the effects of such evolution. If a population
is constantly subject to exogenous environmental change, a steady
state can never be attained, but to the extent that the environment
is itself made up of (or is the effect of) evolving organisms, it makes
sense to ask whether a *general equilibrium* – where all organisms
have attained local maxima relative to each other – can be
realized.[11]

The conditions for the existence and attainability of a general
biological equilibrium could presumably be very complex. Here
I only stress the general idea that evolution may be seen under
the increasingly complex assumptions of a parametric environ-
ment, a strategic (or, in view of I.4, quasi-strategic) environment
and an exogenously changing environment. In the abstract model
used here I deal with the first case only, though I believe that the
conclusions are easily extended to the second case. With the third
case, however, entirely different considerations become relevant.
In particular the phenomenon of *preadaptation* – stemming from
the fact that 'any structure has properties beyond those for which
it was constructed'[12] – permits the attainment of global optima that
in a constant environment would have remained inaccessible. This,
however, is a chance phenomenon only, not a general mechanism
comparable to natural selection.

Enough should have been said by now to explain why the
realization of local maxima in the organic world is not a matter
of course. There is no logical objection to the idea of a world where
the rate of change of the environment relative to the mutation rate
is so high that most organisms most of the time are badly adapted
to each other and to their inorganic environment. In the world
we know, however, the infinitely subtle adaptations found in the
structure and behaviour of organisms are facts that for millennia
have evoked the wonder and (with less justification) the admiration
of naturalists. In many well-documented cases the natural solution

[11]   Winter (1971), p. 258.   [12]   Salthe (1972), p. 9.

to structural and functional problems is strikingly close to the solution that would have been chosen by the engineer or the economist working on the same problem. In some cases animals and men *are* facing the same problems, so that the actual solutions can be compared. As shown by d'Arcy Wentworth Thompson in his classic work *On Growth and Form* as well as by several recent authors,[13] these solutions are often strongly convergent. In recent ecological work[14] nature is seen as an economist rather than as an engineer. Optimal budgeting, linear programming, profit maximization and cost minimization are now as much part of evolutionary theory as of economics.

It is sometimes said that the theory of natural selection is a tautology; that the survival of the fittest means the survival of the survivors. If by fitness we mean genetic fitness, this is true enough. If, however, we understand fitness as ecological fitness, as measured for example by the life span of the organism, the survival of the fittest is turned into an empirical proposition to which there are many important counterexamples. To take but one example, natural selection works so as to produce the optimal sacrifice of parents for offspring. Too much sacrifice for a given offspring may reduce the chance of having more offspring later on, whereas too little sacrifice would damage the 'investment' already made in the offspring. If, on the other hand, the aim was to maximize ecological adaptation, zero sacrifice would clearly be the optimal strategy; indeed the best strategy would be to have no reproduction at all. The confusion between ecological and genetic fitness occurs in various forms in the literature. A rather crass example is found in a recent work on the philosophy of biology, where the author claims that 'the process of childbearing in humans' shows that 'what may be useful to the group (e.g. a species) may have no utility at all for the individual or may even have negative utility'.[15] A more subtle error (or potential for error) underlies the frequent use of the term 'parental investment' in recent evolutionary theory.[16] Whereas economic investment means

---

[13]   Leigh (1971), part I; Frazzetta (1975), ch. 5 and *passim.* On d'Arcy Thompson, see Gould (1976).
[14]   For surveys see Cody (1974) and Rapport and Turner (1977).
[15]   Simon (1971), p. 82.
[16]   Cp. especially Trivers (1972), p. 139, who defines parental investment as '*any investment by the parent in an individual offspring that increases the offspring's chance of surviving (and*

the sacrifice of something (i.e. consumption) now in order to get more *of the same thing* later, parental investment means the sacrifice of something (i.e. ecological fitness) in order to get *more of another thing* (i.e. genetic fitness).

I.3 THE GLOBALLY MAXIMIZING MACHINE

For the purposes of the present argument the crucial features of the locally maximizing machine refer to what it *cannot* do. In particular the machine is incapable of *waiting* and of using *indirect strategies*. These notions are defined as follows. The machine is capable of waiting if it can say No to a favourable mutation in order to be able to say Yes to an even more favourable one later on. Suppose that of a given protein, form B is superior to form A and form C is superior to B, and that because of the structure of the genetic code one-step mutations are possible from A to B and from A to C, but not from B to C. If the mutation to B should occur first, the organism is incapable of saying No. The machine is capable of indirect strategies if it can say Yes to an unfavourable mutation in order to be able later on to say Yes to a very favourable one. Suppose that A is superior to B and C to A, and that one-step mutations can occur from A to B and from B to C, but not from A to C. If B comes about, the organism will not be able to say Yes. As a gradient-climber the organism has its myopic eye fixed to the ground, and is incapable of taking account of what happens behind the next hill. As one author puts it, 'future events or possibilities simply cannot have any effect on the gene pools of organisms'.[17] It is no doubt true that 'even mutations to codons that are completely synonymous with each other may in many instances alter the future mutational possibilities of the genotype',[18] but these possibilities cannot set up a selectional pressure *now* that could give an edge to an apparently neutral mutation over the *status quo*.

By contrast, waiting and the use of indirect strategies are crucial

hence *reproductive success*) *at the cost of the parent's ability to invest in other offspring'.* Investment in the economic sense would rather be 'any investment by the parent in an individual offspring that increases the offspring's chance of having surviving offspring (and hence reproductive success) at the cost of the parent's chance of having surviving offspring'.

[17] Salthe (1972), p. 133.   [18] Stebbins and Lewontin (1972), p. 24.

10   I.3   *The globally maximizing machine*

features of human choice. I suggest that man may indeed be seen as a *globally maximizing machine*, a characterization that goes back to Leibniz, who saw man as being uniquely capable 'reculer pour mieux sauter'.[19] Some examples of global maximization from economic and political life are:

	Waiting	Indirect strategies
Politics	Anti-activism	Anti-reformism
Economics	Patent system	Investment

A few comments may be in order. Investment is perhaps the simplest example of global maximization that requires bypassing a local maximum: one step backwards in order to take two steps forwards. As observed above, parental investment in the animal realm is not really an example of investment in this sense. If, for example, one generation accepted a less-than-maximal number of offspring for the sake of a larger number of grandchildren, this would indeed be investment in the strict sense, but no such idea has ever been entertained in the biological literature. The patent system has the paradoxical feature that 'by slowing down the diffusion of technical progress it ensures that there will be more progress to diffuse'.[20] As Joseph Schumpeter stressed, the maximal exploitation of present possibilities may often be an obstacle to the maximal creation of new possibilities.[21] As suggested by Maurice Meisner, this proposition may also be applied to activism as an

[19] Leibniz (1875–90), III, pp. 346, 578; VII, p. 568. These references actually invoke indirect strategies as a part of the *divine* rationality in the construction of the best of all possible worlds. Similar propositions about *man* also abound, such as the following: 'Les appétitions sont comme la tendance de la pierre qui va le plus droit mais non pas toujours le meilleur chemin vers le centre de la terre, *ne pouvant pas prévoir* qu'elle rencontrera des rochers où elle se brisera, au lieu qu'elle se serait aprochée davantage de son but, si elle avait eu *l'esprit et le moyen de se détourner*...Le bonheur est donc pour ainsi dire un chemin par des plaisirs; et le plaisir n'est qu'un pas et un avancement vers le bonheur, le plus court qui se peut faire *selon les présentes impressions*, mais non pas toujours le meilleur. On peut manquer le vrai chemin, en voulant suivre le plus court, comme la pierre allant droit, peut rencontrer trop tôt des obstacles, qui l'empêchent d'avancer assez vers le centre de la terre. Ce qui fait connaître, que c'est la raison et la volonté qui nous mènent vers le bonheur, mais que le sentiment et l'appétit ne nous portent que vers le plaisir' (V, pp. 175, 182; italics added). Cp. Elster (1975a), ch. VI, for a further discussion.
[20] Robinson (1956), p. 87.   [21] Schumpeter (1954), p. 87.