WILLIAM C. WIMSATT

REDUCTIONISTIC RESEARCH STRATEGIES AND THEIR BIASES IN THE UNITS OF SELECTION CONTROVERSY*

1. MOTIVATING REMARKS ON GENETIC DETERMINISM

"A hen is but an egg's way of making another egg."
Samuel Butler

Butler's satiric comment encapsulates the reductionistic spirit that made Darwinism objectionable to many in his own day, but has fared ever better as a prophetic characterization of the explanatory tenor of modern evolutionary biology. It preceded August Weismann's doctrine of the continuity of the germ plasm advanced in his inaugural lecture in 1883 by some five years. As Weismann's views became one of the anchor points of the modern 'neo-Darwinian' theory of evolution, they led to many modern recapitulations and elaborations of Butler's epigram. Thus Richard Dawkins writes (1976, p. 21):

> Was there to be any end to the gradual improvement in the techniques and artifices used by the replicators to insure their own continuance in the world? . . . They did not die out, for they are past masters of the survival arts. But do not look for them floating loose in the sea . . . . Now they swarm in huge colonies, safe inside gigantic lumbering robots, sealed off from the outside world, communicating with it by torturous indirect routes, manipulating it by indirect control. They are in you and me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. . . . Now they go by the name of genes, and we are their survival machines.

Thus Dawkins announces his intention to " . . . argue that the fundamental unit of selection, and therefore of self-interest, is not the species, nor the group, nor even, strictly, the individual. It is the gene, the unit of heredity" (1976, p. 12). His purple prose gives ample food for worries that the account of evolved structure and behavior in general, and social behavior in particular, will do at best 'simple justice' to its complexities and smack of genetic determinism. But his conclusions are well anchored in the dominant interpretation of the modern 'genetical theory of natural selection', as R. A. Fisher called his theory (Fisher, 1930), and are espoused by many major students of evolutionary biology, as exemplified in the works of G. C. Williams (1966), J. Maynard Smith (1975), and E. O. Wilson (1975). Dawkins has in fact been a

clear and ingenious expositor and elaborator of this view (1976, 1978), despite his often colorful language.

The quote from Dawkins is a direct reflection of the genetic determinism espoused by many and perhaps by most evolutionary biologists today. I take this view to involve two theses, one ontological and the other dynamic about the nature of evolutionary processes:

T1: (Ontological thesis): Genes are the only significant units (or individuals) required for the analysis of evolutionary processes.

T2: (Dynamical thesis): Processes at the genetic level determine (and are the primary and ultimate) explanations for processes at all higher levels.

Genetic determinism has its origins in a misconstrual of the nature of reductionism and of reductive explanation promulgated by most philosophers and by many biologists. A correct view makes it plausible, even inevitable:

(1) that there should be a variety of significant units of selection at various level of organization, thus denying *T1*;

(2) that understanding evolutionary processes requires the invocation and analysis of causal mechanisms and nomic regularities concerning their behavior at each of these levels of organization for the explanation of phenomena at a variety of levels, including that of the individual gene and a number of higher levels, thus denying *T2*;

(3) that sociobiology, properly conceived, should be viewed as the incorporation of an evolutionary perspective into the analysis of processes at these levels (and for most aspects of human social evolution, invoking cultural, rather than biological evolution) rather than the replacement of sociology, psychology, anthropology and the other social sciences by an extended, genetically based ethology of the selfish gene; and

(4) that the apparent success and power of genetic reductionist theories derives from distortions produced by cognitive biases arising from the uncritical application of a variety of reductionistic problem solving heuristics and research strategies.

In what follows I will outline the standard philosophical account of reduction and how it relates to genetic determinism (Section 2); discuss its inadequacy in handling problems of computational complexity (Section 3); show in particular how a claim of *in principle* reduction made by G. C. Williams fails for the simplest extension to a more complex system than that of two alleles at one locus (Section 4); and then discuss at length (Sections 5 through 8) how the problem solving heuristics actually used by reductionistically inclined scientists result in systematic distortions biasing the case against recognizing the need for invoking various higher-level units of selection. A study of these biases suggests recommendations for methodological procedures which should at least partially mitigate their effects, and will in any case serve as a warning to those who must use these heuristic procedures.

2. THE PHILOSOPHERS' VIEW OF REDUCTION

The view I will be discussing in generically based on the model of Nagel (1961) and has been elaborated in somewhat different representative directions by Schaffner (1967, 1969, 1974, 1976), Ruse (1974, 1976), and Causey (1972*a*, 1972*b*, 1976). It has been widely criticized by a number of authors, including Hull (1973, 1974, 1976), Nickles (1973, 1976), Dresden (1974), Darden and Maull (1976), Maull (1977), Bantz (this volume), Bogaard (1979), and myself (1974, 1976*a*, 1976*b*, 1978). My (1978) is an extensive review of the literature, but probably the best self-contained systematic critical analysis of these issues is to be found in McCauley (1979, Chapters 4 and 5). Schaffner's own most recent work (1979), while not explicitly discussing reduction, seems to me to be a powerful and productive move away from his earlier position, in this direction. I mention these sources to indicate where extensive discussion of relevant issues can be found. Most of this discussion will be presupposed here.

The traditional view of reduction holds that it is the

(i) *in principle*

(ii) *deducibility* of upper-level entities, properties, theories, and laws

(iii) in terms of the properties, laws, and relations *of any degree of complexity* of entities at the lower level.

I have emphasized in this characterization the three clauses which have caused the greatest problems for this traditional view. This view has two corollaries:

C1: Upper level entities are thus shown to be 'nothing more than' collections of lower level entities (and their relations).

C2: Upper level laws and causal relations are illusory or are shown to be 'nothing more than' a shorthand for and to be determined by lower level laws and causal relations.

These corollaries have a direct relation to the ontological and dynamical theses of genetic determinism, for *C1* and *C2* respectively provide the reasons for holding *T1* and *T2* to be true, if the relation of the various higher-level units of selection and phenomena to those at the genetic level is one of reduction, as it has been traditionally construed. *T1* is true *because* then the various higher-level units of selection are 'nothing more than' collections of genes (and their relations) in a fashion demonstrated by the deductive or definitional relations between terms, and similarly for *T2* and *C2*.

A major problem for applying the Nagelian model of reduction, one recognized by many of its defenders (see, *e.g.*, Schaffner, 1974), is that it appears not to fit the practice of reductionistically inclined scientists. As various writers have observed, if the Nagel model of reduction as a kind of deduction or its extensions is accepted as an adequate model of reduction, there may not *be* any cases of reduction in science (see Schaffner, 1974; Hull, 1973, 1976). A standard retreat in the face of this problem has been from claiming deducibility in practice, to defending a claim of deducibility or analyzability *in principle*. For this reason, I have added the 'in principle' qualifier explicitly in my characterization of reduction, since at least this modification is required to describe the actual practice of scientists.

Further problems have arisen with the characterization of the analysis of the upper level in lower level terms as a kind of deduction. Dresden (1974), Bogaard (1979), Bantz (this volume), Sklar (1973, 1976), and others have pointed to the role of approximations, which prevent this 'derivation' from being characterized as a deduction. I have emphasized (1976*b*, pp. 685–689) how the interpretation and elaboration of the implicit *ceteris paribus* clause in purported deductions makes any supposed 'translation' context-dependent in a way that undercuts the usefulness of the deductive model and falsifies corollaries *C1* and *C2*. Furthermore, not only are there problems with filling out the *ceteris paribus* clause, but, rather anomalously on this model, scientists appear to have no interest in trying to do so.

I have argued elsewhere that the 'in principle' deducibility, analyzability, or translatability is best seen not as the primary structure, focus, or thesis of reductionism but as a derivative corollary of the use of identificatory hypotheses in reductive explanation. If an upper level entity, phenomenon, *etc.* is *identified* with a lower-level complex of entities, properties, and relations, then Leibniz's law tells us that any property of one is a property of the other. Thus *Leibniz's Law* tells us that if the purported identity holds, any upper level thing must *in principle* be analyzable in lower-level terms. These 'in principle' claims thus become an important heuristic method for moving

from an identity claim to specific hypotheses about heretofore unmatched properties at the upper and lower levels (see Wimsatt, 1976*a*, pp. 225–237, and 1976*b*, pp. 697–701). This heuristic use of identity claims and *in principle* arguments provides further support for the view advocated here: that *the power, limitations, and character of reductionistic approaches in science is better analyzed in terms of the reductionistic research strategies one is led to adopt than in terms of idealized deductive accounts and ontological theses derived from them*. For more on this use of identity claims, see my (1976*a*; 1976*b*, Section 3; and 1978). It will not be further elaborated here.

The use of simpler models and approximations in reductionistic modelling produces a gap between promise and performance that has interesting consequences. The metaphysical position that the reductionist defends holds that a reductionistic analysis of upper-level phenomena must exist in terms of lower level entities, properties, and relations *of some degree of complexity* – preferably in terms of monadic properties; but if not these, then at least in terms of *some* (possibly complex and relational) properties of the lower-level entities. This is another formulation of the claim of *in principle* deducibility of reducibility that I have argued is a corollary of the use of identity claims and Leibniz's Law.

The holist, as anti-reductionist, is taken normally as denying this metaphysical claim, and thus to be holding the equally metaphysical (and to most people, radically implausible) claim that no analysis of whatever complexity in lower-level terms could be adequate. But, despite appearances, the in principle claim of the reductionist is seldom in dispute. *In the cases I know in population biology, in neurophysiology, and in the history of genetics, the issue between scientists who are reductionists and holists is not over the in principle possibility of an analysis in lower level terms but on the complexity and scope of the properties and analyses required*. The more holistically inclined scientists usually argue that higher-order relational properties of the lower-level entities are required, and the reductionists argue that a given simple, lower level model (often one using only monadic properties) is adequate. To the extent that this is true, the portrayal of the dispute between reductionist and holist as over the *in principle* claim (a portrayal favored by most philosophers, and by many scientists) is seriously in error and turns a usually serious, comprehensible, and important empirical dispute into a usually one-sided and poorly motivated metaphysical one.

This reading of the dispute might seem to have the apparent disadvantage of dissolving it, for both holists and reductionists now appear to be species of reductionist – "complex" reductionists or "simple" reductionists respectively.

But this species of complex reductionist is still recognizably a holist. (See Wimsatt, 1978.) Complex relations relate several to many lower-level entities, and require the recognition of these complexes as entities. Further, if the relationships overlap in their relata, these higher-level entities become tied together in terms of still higher level systems in ways suggested by the discussion of descriptive and interactional complexity in (Wimsatt, 1974). The further presence of many-one mappings between lower and higher-level state descriptions, required even by the existence of recognizable stable higher-level phenomena generates a kind of autonomy and independence of the dynamics and the explanations at the higher level from detailed lower level specifications and laws. (See Wimsatt, 1979, Sections 4 and 5, for the discussions of 'sufficient parameters' and 'robustness'; Wimsatt, 1976*a*, pp. 248–251 on 'explanatory primacy'; and Wimsatt, 1976*b*, section 6, pp. 689–692, and the appendix, pp. 701–704, for the discussion of the 'screening off' relation of Salmon.)

The net effect of these considerations is that the holist can get the significance and autonomy of upper level entities, laws, and phenomena which he desires while accepting a kind of in principle (but 'complex') reductionism. The arguments in what follows for the significance of higher-level units of selection are to be interpreted as espousing this kind of holism.

3. THE PROBLEM OF COMPUTATIONAL COMPLEXITY AND THE USE OF REDUCTIONISTIC RESEARCH HEURISTICS

I claimed in the last section that the claims of 'in principle' deducibility or translatability are best seen as corollaries of Leibniz's Law, and thus as consequences of the use of compositional identities in reductive explanations. However, there is much to be learned by looking at the standard explication of these claims. Most of these seem to suppose that a claim of *in principle* translatability is to be explicated in terms of effective computability or, mirroring Laplace's definition of a deterministic system, as a translation which could be produced by a sufficiently powerful computer which was given a total state description of the micro-level of the appropriate system, together with all of the micro-level laws which applied to that system. Richard Boyd (1972) has given a brilliant criticism of this possibility in general. I wish here to make only some more pragmatic criticisms (see also my 1978).

First of all, it is unclear how a practicing scientist could make use of results concerning the effective computability of any system he is studying, since all of these results would presuppose a total knowledge of the system,

which he does not possess, and a theory of that system organized in a fashion unlike any of the theories which he knows. It is all right to *talk* about writing the Schrödinger wave equation for a particular organism, but in fact physical and quantum chemists don't even do it for chemical bonding in simple molecules (see Bantz, this volume; Bogaard, 1979; and Dresden, 1974). Instead they use simple heuristic approximations even for these far simpler cases.

But suppose that one could. Does this mean that we would, or even could study systems in this way? It does not. This can be seen by studying the game of chess. Chess is a totally deterministic game. At each stage, the possible moves and their outcomes are exhaustively specifiable – and indeed relatively straightforwardly specifiable. This means that if we specify in advance how many moves we wish to allow in the game, we can in principle write down a branching tree, beginning with the initial state of the chessboard and ending with branches of nodes corresponding to all possible games of that many moves or less. (Some of the games may already have terminated in fewer moves.)

There are twenty possible opening moves (two for each of eight pawns and two knights). Suppose that this number of alternatives continues throughout the game on the average, as a geometrical mean. (This is almost certainly an underestimate.) And suppose that we wish to consider games of 100 moves (fifty pairs of moves). Then we are considering on the order of 20^{100} possible games. This is a large number, but clearly it is a task which is effectively computable. But that is not much consolation. Consider the size of the task: $20^{100} = 2^{100} \times 10^{100}$. Since $2^{10} \approx 10^3$, then $20^{100} \approx 10^{130}$. Now for some other relevant numbers. There are about 10^{79} elementary particles in the universe. There have been about 10^{19} seconds since the big bang. And the shortest known time for a physical event is on the order of 10^{-24} second, the time it takes for light to traverse the diameter of an atomic nucleus. Putting this all together, we arrive at an upper estimate of the number of events in the universe to date of about 10^{122}. Then this task, a trivial one for a universal Turing machine, is nonetheless *not* doable by the most universal computer we could imagine – the universe as a computer! It would fall 8 orders of magnitude short of having had enough *actual* states (as opposed to possible ones) to represent all of these games, and we have not even raised the questions of how these games could be mapped into the states in a usable manner and how rapidly the different parts of a computer spanning 10^{10} light years will be able to communicate with one another! Clearly, this effectively computable (and therefore, to many logicians and mathematicians, *in principle* possible) task is, ***physically speaking, in principle impossible***

This kind of observation led Herbert Simon and others since to look for other kinds of models than the exhaustive, brute force algorithmic approach for human problem solving. First, for decision making (Simon, 1957), then for proving theorems (Newell, Shaw, and Simon, 1958), and subsequently for other problems (Newell and Simon, 1961; Simon, 1966*a*, 1966*b*, 1969, 1973), Simon espoused a 'principle of bounded rationality' (Simon, 1957, pp. 198–199) which asserts that we are generally faced with problems of such complexity that we cannot solve them exactly, and therefore, if we are to get any solutions at all we must do so by introducing various simplifying and approximative techniques. Thus was born the idea of a heuristic.[1] As I use that notion here, I take a heuristic procedure to have three important properties:

(1) By contrast with an algorithmic procedure, the correct application of a heuristic procedure does not guarantee a solution; and if it produces a solution, does not guarantee that the solution is correct. Thus valid deduction from true premises is not a heuristic procedure. Most or all inductive procedures are, however (see Shimony, 1970).

(2) The expected time, effort, and computational complexity of producing a solution using a heuristic procedure is appreciably less than that expected using an algorithmic procedure. This is indeed the reason why heuristics are used. They are a 'cost-effective' way of producing a solution, and often the only physically possible way.

(3) The failures and errors produced using a heuristic are not random, but systematic. I conjecture that *any heuristic, once we understand how it works, can be made to fail*. That is, given this knowledge of the heuristic procedure, we can construct classes of problems for which it will always fail to produce an answer, or for which it will always produce the *wrong* answer. This property of systematic production of wrong answers will be called the *bias(es)* of the heuristic.

Not only can we work forward from an understanding of a heuristic to predict its biases, but we can also work backwards, hypothetically, from the observation of systematic biases as data to conjecture as to the heuristic which produced them; and if we can get independent evidence as to the nature of the heuristics, we can propose a well-founded theory of the structure of our heuristic reasoning in these areas. This was elegantly done for the

first time by Tversky and Kahneman (1974), in their analysis of fallacies of probabilistic reasoning and the cognitive heuristics which produce them. To my mind, Simon's work and that of Tversky and Kahneman have opened up a whole new set of questions, a new area of investigation of pragmatic inference in science, which should revolutionize our discipline in the next decade, and increasing numbers of workers are moving in this direction. (See, for example, the papers of Schaffner, Darden, and Bantz in this volume.)

The notion of a heuristic has far greater implications than can be explored in this paper. In addition to its centrality in human problem solving, it is a pivotal concept in evolutionary biology and in evolutionary epistemology. It is a central concept in evolutionary biology because any biological adaptation meets the conditions given for a heuristic procedure. First, it is a commonplace among evolutionary biologists that adaptations, even when functioning properly, do not guarantee survival and production of offspring. Secondly, they are, however, cost-effective ways of contributing to this end. Finally, any adaptation has systematically specifiable conditions, derivable through an understanding of the adaptation, under which its employment will actually *decrease* the fitness of the organism employing it, by causing the organism to do what is, under those conditions, the wrong thing for its survival and reproduction. (This, of course, seldom happens in the organism's 'normal' environment, or the adaptation would become maladaptive and be selected against.) This fact is indeed systematically exploited in the functional analysis of organic adaptations. It is a truism of functional inference that learning the conditions under which a system malfunctions, and how it malfunctions under those conditions, is a powerful tool for determining how it functions normally and the conditions under which it was designed to function. (See, *e.g.*, Gregory, 1967; Lorenz, 1965; Valenstein, 1973; and Glassman, 1978, for illuminating discussion of the problems, techniques, and fallacies of functional inference under a variety of circumstances.)

The notion of a heuristic is central to evolutionary epistemology, because Campbell's notion of a 'vicarious selector' (1974, 1977), which is central to his conception of a hierarchy of adaptive and selective processes spanning subcognitive, cognitive, and social levels, is that of a heuristic procedure. A vicarious selector for Campbell is a (1) substitute (2) less costly selection procedure acting to optimize some index which is only contingently connected with the index optimized by the selection process it is substituting for. This contingent connection allows for the possibility – indeed the inevitability – of systematic error when the conditions for the contingent concilience of the substitute and primary indices are not met. An important ramification of

Campbell's idea of a vicarious selector is the possibility that one heuristic may substitute for another (rather than for an algorithmic procedure) under restricted sets of conditions, and that this process may be repeated, producing a nested hierarchy of heuristics. I believe that this is an appropriate model for describing the nested or sequential structure of many approximation techniques, limiting operations, and the families of progressively more realistic models found widely in 'progressive research programs', as exemplified in the development of 19th century kinetic theory, early 20th century genetics, and in several areas of modern population genetics and evolutionary ecology. (On this last, see, *e.g.*, Roughgarden, 1979.)

The ultimate end of this paper is to discuss some of the heuristic used in reductionistic modelling and to show how their systematic biases have given illegitimate support to a reductionistic vision of evolutionary processes culminating in genetic determinism. But before I do this, it is necessary to show that, how, and why a brute force, quasi-algorithmic, reductionistic approach cannot work in evolutionary biology and population genetics, for just such an approach has been suggested by G. C. Williams.

4. WILLIAMS'S 'IN PRINCIPLE' REDUCTIONISM AND THE CASE OF TWO LOCI

A problem-solving heuristic which Simon (1966) has called 'factoring into subproblems' appears in a variety of guises in reductionistic modelling. Simon illustrates the heuristic and its advantages using the problem of finding the right combination for a combination lock. Imagine a bicycle lock with ten wheels of ten positions each. If there is only one combination which will work, one would expect to look through about half of the possible 10^{10} combinations on the average before finding it. On the other hand, suppose that the lock is a cheap or defective one for which one can tell individually for each wheel when it is in the right position. Then an average of 5 tries on each wheel, for a total of 50 tries would be expected to find the right combination. The advantage that accrues from being able to break the problem down into subproblems, being able to find out parts of the combination, rather than having to solve the whole problem at once, is given by the ratio of the number of alternatives which must be inspected. This is, in this case, $(5 \times 10^9)/(5 \times 10) = 10^8$.

Similar advantages accrue for similar combinatorial reasons if problems of evolutionary dynamics can be treated in terms of the frequencies of individual alleles, with no epistatic interactions and no probabilistic associations between

alleles at different loci due to linkage or assortative mating rather than in terms of the gametic or zygotic genotype frequencies required if these assumptions do not hold. Here the simplification occurs in the number of dimensions in the phase space required to adequately describe and predict evolutionary changes, and, correlatively, in the number of state variables in the equations required to describe and predict the dynamics of evolutionary change. Table 1, derived and extended from Table 56 of Lewontin (1974, p. 283), summarizes the dimensionality of the problem under different simplifying assumptions. It is worth noting that if *no* simplifying assumptions are

TABLE 1

Sufficient dimensionality required for the prediction of evolution of a single locus with a alleles where there are n segregating loci in the system

Level of Description:		Zygotic Classes	Gametic Classes	Allele Frequencies	Allele Frequencies
Dimensionality:		$\frac{a^n(a^n+1)}{2}-1$	a^n-1	$n(a-1)$	$(a-1)$
Assumptions:		none	1	1, 2	1, 2, 3
n:	a:				
2	2	9	3	2	1
3	2	35	7	3	1
3	3	377	26	6	2
5	2	527	31	5	1
10	2	524799	1023	10	1
32	2	9.22×10^{18}	4.29×10^9	32	1

Assumptions:
(1) random union of gametes (no sex linkage, no assortative mating)
(2) random statistical association of genes at different loci (linkage equilibrium).
(3) no epistatic interaction (inter-locus effects are totally additive).

(Table is adapted and extended from Table 56 of Lewontin, 1974, p. 283.)

made, even the simplest multi-locus case of two alleles at each of two loci is analytically intractable. This should not be surprising: the problem of dimensionality nine (there are nine possible genotypes, with independently specifiable fitness parameters) is already more complicated than the three-body problem of classical mechanics. Like the three-body problem, it has been

solved for a variety of special case (see Roughgarden, 1979, Chapter 8, pp. 111–133) but has not been solved in general.

In the light of this, G. C. Williams makes a claim which gives substantial hope, for it appears to promise that the problem can be treated as one of the lowest dimensionality. His view is that since the operation of any higher-level selection processes can be mathematically expressed as resulting from the operation of selection coefficients acting independently at each locus to change the frequency of individual alleles or genes, there is no need to postulate the existence of any higher-level units of selection or selection forces. This view will look most familiar to philosophers, since it bears the strongest resemblance to traditional philosophical accounts of theory reduction. Williams expresses it as follows:

> Obviously it is unrealistic to believe that a gene actually exists in its own world with no complications other than abstract selection coefficients and mutation rates. The unity of the genotype and the functional subordination of the individual genes to each other and to their surroundings would seem, at first sight, to invalidate the one-locus model of natural selection. Actually these considerations do not bear on the basic postulates of the theory. No matter how functionally dependent a gene may be, and no matter how complicated its interactions with other genes and environmental factors, it must always be true that a given gene substitution will have an arithmetic mean effect on fitness in any population. One allele can always be regarded as having a certain selection coefficient relative to another at the same locus at any given point in time. Such coefficients are numbers that can be treated algebraically, and conclusions inferred from one locus can be iterated over all loci. Adaptation can thus be attributed to the effect of selection acting independently at each locus. (Williams, 1966, pp. 56–57)

Williams goes on, in the next two pages, to illustrate how this algebraic manipulation can be accomplished in a simplified genetic environment of two alleles at each of two loci, and we are to imagine the extrapolation to cases of many alleles at many loci. Complicated it would be, but *in principle*, of course (we are told), it could be done, "by iterating over all loci."

This claim might appear to involve another variant of the 'factoring into subproblems' heuristic which Simon has studied and written upon at length, and which he has called 'the hypothesis of near decomposability' (see Simon, 1969, pp. 99ff, in Ando *et al.*, 1963, and further references given there; and also Wimsatt, 1974, for further discussion).

The hypothesis of near decomposability involves the assumption that a complex system can be decomposed into a set of subsystems such that all strong interactions are contained within subsystems' boundaries, and interactions between variables or entities in different subsystems are appreciably weaker than those relating variables or entities in the same subsystem. In this

case, an approximation to the behavior of the system, in the short run, can be gotten by ignoring the intersystemic interactions and analyzing each subsystem as if it were isolated, studying only internal variables in their common approach to equilibrium. Its behavior in the long run can be approximated by ignoring the intra-systemic interactions of the subsystems (assuming that there is intra-systemic equilibrium), representing each subsystem by a single index, and considering the equilibration of the various subsystems with one another as a system involving the interaction only of these lumped index variables.

There are thus two different approximations involved in studying the short-run and the long run behavior of the system. Each substantially reduces the complexity of the problem, if the assumptions allowing the approximation are justified.

Indeed the hypothesis of near decomposability *is* used in this way in a number of multi-locus models – in particular when it is assumed: (1) that the system starts at or near linkage equilibrium. (This is the condition when all genotypes occur at frequencies given by the products of the frequencies of their constituent genes, a condition equivalent to the assumption of a multi-locus Hardy-Weinberg equilibrium.) (2) That selection between genotypes is relatively weak (a condition that guarantees that the population never deviates far from the multi-locus Hardy-Weinberg equilibrium). Indeed, these assumptions (as well as that of random association of gametes, implying no assortative mating) are made in the original model (Lewontin and White, 1960) for which the two-locus fitness surfaces, which provide below a counterexample to Williams's claim, originally were derived. Under these conditions, recombination can be neglected as a significant contributor to genotype frequencies, and the dynamics can be treated as if they are affected by segregation and selection only. This is equivalent to using the 'long range' approximation in studying the behavior of a nearly decomposable system, since if the system is far from linkage-equilibrium, recombination may be a far greater contributor to genotype frequency of some genotypes than either segregation or selection, and thus behaves like an intra-systemic 'strong' interaction which goes relatively rapidly to equilibrium.[2] The observation that under some conditions there can be permanent and substantial linkage disequilibrium (see Lewontin, 1974; Roughgarden, 1979; and also Maynard Smith, 1978, Chapter 5) is equivalent then to saying that the system cannot be treated as nearly decomposable.

In fact, however, Williams's claim in the above quote appears to be far stronger than a near decomposability claim and is not made on the basis of

these assumptions about linkage equilibrium and random assortment of gametes produced by random mating. He claims that the problem can be solved one locus at a time and then extended to a global solution by "iterating over all loci." His claim is thus not that the genetic system is nearly decomposable, but that it is simply decomposable, like the simpler of Simon's two locks. Without all of the qualifications in Lewontin's table, this claim is simply incorrect, and can be shown to be so for the simplest case involving more than one locus – that of two alleles at each of two loci. The reasons for the failure of Williams's claim can best be seen after a discussion of this case.

A claim that evolutionary processes can be analyzed in the manner Williams suggests, as being of the lowest possible dimensionality, involves at least the claim that a deterministic theory of the change of gene frequencies at a given locus can be constructed using only the frequencies of the alternative alleles of that locus. In the simplest case of two alleles at one locus, this involves saying that it is a function only of the frequency of a single gene since if q is the frequency of gene a, then $1-q$ must be the frequency of the other gene, A because there are not other genes at that locus. (It is a function also of the fitnesses, W_{11}, W_{12} and W_{22} of the genotypes AA, Aa, and aa, but these are assumed to be constant parameters of the system in this discussion.)

Consider Figure 1 as a graph of gene frequency from different initial points (.05 for the bottom curve, .95 for the top curve) as it changes in successive generations. If this were the graph of an actual case (Lewontin describes it as of a "hypothetical laboratory population") it would falsify Williams's claim. Why? Consider the topmost curve. At all points between the initial high value of gene frequency of .95 and the minimum value (of about.7, reached in generation 4) a population which is decreasing in gene frequency at that value (between generations 0 and 4) is later increasing in gene frequency at that value (in generations 5 and later). *But if gene frequency can either increase or decrease from a given value, then gene frequency* (of that gene or its allele) *alone is not an adequate basis for a deterministic theory of evolutionary change.*

Williams gets into trouble at this point because his claim is neither a theory of evolution in terms of gene frequencies, nor even a schematic description of the form of such a theory. His statement that " . . . it must always be true that a given gene substitution will have an arithmetic mean effect on fitness in any population" (1966, p. 56) suggests the following procedure for evaluating this effect on fitness. Imagine a gigantic (non-interventive) DNA sequencer that, given a population, will determine all of the genes in that population and their frequencies. Perform this genetic census at two points in time – or

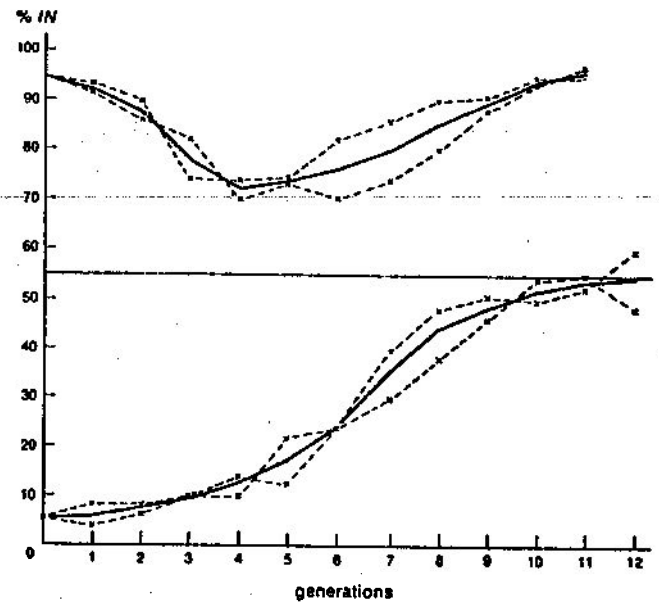

Fig. 1. The frequency of an inversion, IN, in hypothetical laboratory populations. The heavy lines represent the average behavior of replicates, while the X's represent individual data points. (Reprinted from Lewontin, 1974, Figure 23, p. 274, with permission of Columbia University Press.)

perhaps in each generation. For each gene, its frequency in the interval will either increase (in which case it is being selected for), decrease (it is selected against), or remain constant (it is neutral).

This data can then be used in one of two ways. It all can be used to describe the evolutionary trajectory of the population in its phase space. But then this is not a *theory* of evolutionary change but a description. The fitnesses, W_{11}, W_{12}, and W_{22} inferred from this are merely biological redescriptions of what is happening in successive generations[3] and may undergo arbitrary changes as the 'curve-fitting' parameters that they are. Or the changes observed in one generation may be used to estimate fitness values which are then used to predict future changes. This is more of a process of trend extrapolation using an assumed model rather than a theory itself, but it is at least not totally tautological. To have a predictive tool or theory then, Williams must intend his remarks to describe a process of trend extrapolation.

But here is where the trouble arises. The graph of Figure 1 indicates that local estimates of fitness values *cannot* be used in this way to extrapolate evolutionary trends. After all, gene *a* is apparently being selected *against* in generations 0–4, but subsequently it must be being selected *for*, as it frequency is then *increasing*. To put it more generally, local estimates of fitness or selective value are not valid globally, for other values of the frequency of that and other genes.

The reason why this is not the case becomes apparent in Figure 2. Indeed, Lewontin's hypothetical laboratory population of Figure 1 was not hypothetical at all, but a description of changes that would be expected in a field population of the grasshopper, *Moraba scurra*, whose mean Darwinian fitness

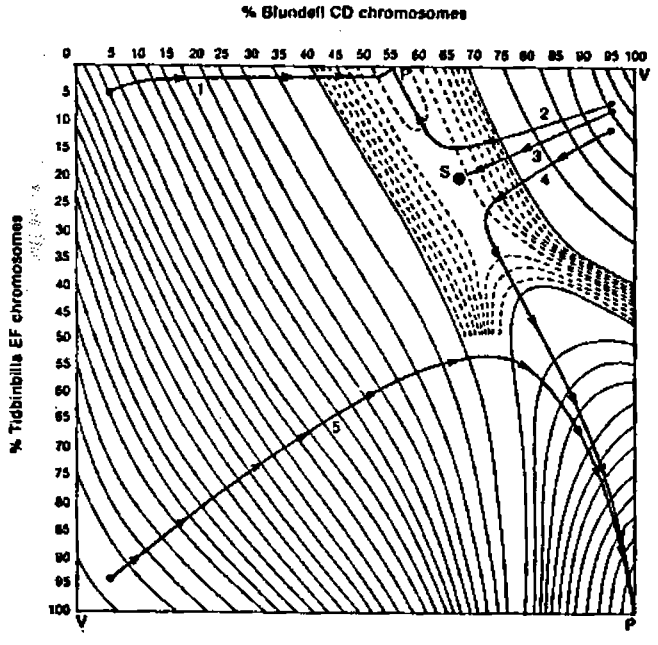

Fig. 2. Projected changes in the frequency of two polymorphic inversion systems in *Moraba scurra* from different initial compositions, based of fitness estimates from nature. The trajectories, shown by arrow-marked lines, are calculated by the solution of differential equations of gene frequency change. Lines crossing the trajectories are contours of equal mean population fitness, $\bar{W}$. (Reprinted from Lewontin, 1974, Figure 24, p. 280, with permission of Columbia University Press.)

$\bar{W}$, is given as a function of the frequency of two alleles of each of two loci in the adaptive topography of Figure 2. An adaptive topography is a plot of contours of equal mean population fitness, $\bar{W}$, as a function of the gene frequencies (in this case, at two loci). Since in many simpler models (particularly where the genotypic fitnesses are constant) a population will tend to evolve in directions of increasing $\bar{W}$, the adaptive topography gives a visual means of making qualitative predictions about the direction and relative rates of local evolutionary change.

Indeed, the lower curve of Figure 1 is just trajectory 1 of Figure 2, and the upper (problematic) curve of Figure 1 is trajectory 4 of Figure 2. And the results that appeared as indeterministic in terms of the trajectory of gene frequencies of a single locus are seen to be deterministic once the frequencies of each of two loci are specified. Thus, the initial points of trajectories 1 and 4 should not have been specified as .05 and .95, the frequency of a gene at the first locus, as is implicit in Figure 1, but as (.05, .05) and as (.95, .12), the frequencies of the genes at both loci as in Figure 2. Note also that trajectory 2, with initial point (.95, .07) – in which the frequency at the first locus is the same as for trajectory 4, but that at the second locus is different – also shows a violation of the deterministic assumption if only the first locus frequency is looked at when the two trajectories are compared. Trajectory 2 is not by itself evidence of a violation, as is trajectory 4, however.

It is quite clear from this adaptive topography that what will happen in evolution is a function of the joint values of gene frequency at two loci, and no set of measurements or extrapolations looking at frequencies of just one locus at a time can provide an adequate basis for prediction. This is true in this case because of epistatic interactions between loci, which is a sufficient condition for having to go to a phase space of greater dimensionality for prediction.

Williams's proposal, then, fails in this case, in 3 ways:

(1) It does not result in a deterministic theory of evolutionary change in terms of the gene frequencies of individual loci which can be "iterated over all loci" to produce a global solution.

(2) It fails to do so because epistatic interactions among loci prevent local estimates of fitness at single loci from being projectable or extrapolatable *if gene frequencies at other loci are free to change simultaneously*. (The appearance of projectability usually arises when estimates are only done locally under conditions in which there *is* no change or no significant change at other loci. But this

cannot be assumed in general.) Williams in effect errs by assuming that single locus fitnesses are independent of context, when in fact they are functions of the context of other loci. Illegitimate assumptions of context-independence are a frequent error in reductionistic analyses. See Wimsatt, 1976, p. 688 and 1980 for further discussion.

(3) In fact, Lewontin's data on *Moraba scurra* represents not gene frequency changes at single loci, but the frequencies of chromosomal inversions involving *many* loci. For reasons which I will not detail here, inversions can often act as units of selection, and Lewontin has devoted much of the earlier portions of his book to arguing that Williams's aim of measuring the fitness effects of single gene substitutions is bedevilled with a host of practical and theoretical problems. So Lewontin's one chromosome example of Figure 1 and the two chromosome counterexample of Figure 2 are already at a higher level of organization than that supposed by Williams's single locus genetic reductionism.

What goes for two loci or chromosomes, goes as well for many. In this light, Williams's remarks suggesting genetic reductionism are better seen as having more import as a kind of genetic bookkeeping than as promising a reductionistic theory of evolutionary change in terms of gene frequencies. The latter is a tempting mirage which vanishes upon closer inspection of the complexities and heuristics of the actual theory.

5. A GENERAL CLASS OF REDUCTIONISTIC RESEARCH HEURISTICS AND THEIR BIASES

The kind of mistake implicit in the reductionistic argument discussed in the preceding section seems so straightforward, yet is so pervasive and is made by so many leading practitioners of the discipline that it cries out for a deeper explanation. Any engineer knows that systematic failures in a mechanism indicate a design problem, which he then tries to locate, and eliminate. This procedure is so important (and so often ignored in the traditional education of engineers) that at least one company holds seminars for the engineers of its various divisions to teach approaches and methods for doing it, as described in Moss (1979). The heuristics in our reasoning processes have similar possibilities for systematic error, and we should similarly try to analyze these failures to get an understanding of where they are likely to occur and, where possible,

eliminate or moderate their effects through redesign. If redesign (through teaching different or modified heuristics) is impossible or impractical, we can at least, through an understanding of the causes of failure, be warned when they are likely to lead us into error, so that our troubleshooting efforts may be concentrated there.

I will describe here a general class of heuristics and their biases which have a common origin in the nature of reductionistic analyses, thus providing the appropriate warnings. After defining the notion of unit of selection, I will describe some of the results of Wade's review of the models of group selection (1978) which nicely illustrates several of these biases with systematic failures in the literature. I will then discuss a design modification in our heuristic which should serve to eliminate or at least moderate the effects of several of these biases.

(1) Any analysis of a system presupposes a division of the world, however tentatively, into the system being studied and its environment. This division may be made on grounds of interest, which will in turn often be determined by judgments of the scope of one's field (a molecular geneticist is unlikely, at least initially, to consider social forces as part of the subject matter of his discipline), other jurisdictional criteria, and probably most frequently, intuitive judgments about the natural chunks and boundaries in his area.[4] Judgments of what can be manipulated relatively independently of 'outside' forces are likely to enter into any of these, and this in turn implies judgments of near decomposability or near isolatability in the individuation of systems.

(2) A reductionist adds to this a further consideration: by his description as a reductionist, he is interested in understanding the behavior of his system in terms of the interaction of its parts. This means that his *interest* at least (though not necessarily his scope of investigation) will be focussed on the entities and interrelations between them *internal* to the system he is studying.

(3) The third and last constraint is to recognize the practical impossibility of generating an exhaustive, quasi-algorithmic, or exact analysis of the behavior of the system in its environment. This is an application of Simon's 'principle of bounded rationality' discussed earlier. So the reductionist must start simplifying. In general, simplifying assumptions will have to be made everywhere, but given his interest in studying relations *internal* to the system, he will tend to order his list of economic priorities so as to simplify first and more extremely in his description observation, control, and analysis of the environment than in the system he is studying. After all, simplifications internal to the system face the danger of simplifying out of existence the very phenomena and mechanisms he wishes to study.

This fact alone, derived just from these three very general assumptions, is sufficient to generate and explain a wealth of heuristics and their attendant biases arising in the reductionistic analysis of systems. These heuristics and biases can be classified roughly as biases of conceptualization, biases of model building and theory construction, and as biases of observation and experimental design, though any rigid classification would fail because of the interdependence and intercalation of these activities in the course of a scientific investigation.

I will here describe them only relatively cursorily, leaving their further elaboration for other occasions. (I have already discussed item 2 in 1976*a*, pp. 244–245, and in 1978. Extensive discussions of items 1, 3, and 4 are forthcoming in Wimsatt, 1980. Wade's work, and the work I have done so far in population biology relates most strongly to items 4, 5, and 6, though 1, 7, 8, and 9 are also implicated.) This is possible because even a statement of the heuristic naturally suggests the pervasiveness of their use and multiplicity of their possible effects, and it is in any case necessitated by space limitations.

These heuristics and/or biases are as follows:

A. *Conceptualization*:

1. *Descriptive Localization*. Describe a relational property as if it were monadic, or a lower order relational property; thus, *e.g.*, fitness as a property of phenotype (or even of genes) rather than phenotype-environmental relation.
2. *Meaning Reductionism*. Assume lower level redescriptions to change meanings of scientific terms; higher level redescriptions not. Result: philosophers (who view themselves as concerned with meaning relations) are inclined to a reductionistic bias.
3. *Interface Determinism*. Assume that all that counts in analyzing the nature and behavior of a system is what comes or goes across the system-environment interface. This has two versions: (a) Black-box behaviorism: all that matters about a system is how it responds to given inputs. (b) Black-world perspectivalism: all that matters about the environment is what comes in across the system boundaries and how it responds to system inputs. Either can introduce reductionistic biases when conjoined with the assumption of white box analysis . . . that the order of study is from a system, with its input-output relations, to its subsystems, with theirs, and so on. The analysis of functional properties in particular, is rendered incoherent and impossible by these assumptions.

B. *Model Building and Theory Construction*:

4. *Modelling Localization*. Look for an intrasystemic mechanism to explain a systemic property, rather than an intersystemic one. Structural properties are regarded as more important than functional ones, and mechanisms as more important than context.
5. *Simplification*. In reductionistic model building, simplify environment before simplifying system. This strategy often legislates higher level systems out of existence or leaves no way of describing systemic phenomena appropriately.
6. *Generalization*. When starting out to improve a simple model of system, environment: focus on generalizing or elaborating the internal structure at the cost of ignoring generalizations or elaborations of the external structure.
   Corollary. If the model doesn't work, it must be because of simplifications in description of internal structure, not because of simplified descriptions of external structure.

C. *Observation and Experimental Design*:

7. *Observation*. Reductionist will tend not to monitor environmental variables, and thus will often tend not to record data necessary to detect interactional or larger scale patterns.
8. *Control*. Reductionist will tend to keep environmental variables constant, and will thus often tend to miss dependencies of system variables on them. ('*Ceteris paribus*' is viewed as a qualifier on environmental variables.)
9. *Testing*. Make sure that a theory works out only locally (or only in the laboratory) rather than testing it in appropriate natural environments, or doing appropriate robustness analyses to suggest what are important environmental variables and/or parameter ranges.

These heuristics and their biases can be particularly powerful for two reasons: (1) There is, on the face of it, no way to correct for their effects – at best not in the most obvious way by producing the exact and general analysis of the behavior of the system in its environment to use as a check against the models produced. That may be all right in theory, but it won't work out in practice, as I have tried to suggest in Section 3. Nonetheless, there are at least two possible corrective measures which will be discussed later. The first is robustness analysis – a term and procedure first suggested

by Richard Levins in his (1966). The second, which I will call 'multi-level reductionistic analysis' involves using these heuristics simultaneously at more than one level of organization -- a procedure which allows discovery of errors and their correction in at least some circumstances, and which in fact implicitly followed as a species of 'means-end analysis' (see Simon, 1966) in the construction of interlevel theories involving compositional identities (see Wimsatt, 1976*a*, pp. 230–237, and 1976*b*, Section 8).

(2) Secondly, it should be clear that *these heuristics are mutually supporting*, not only in their effective use in structuring and in solving problems, but also *in reinforcing, multiplying, and above all, in hiding the effects of their respective biases*. This effect of bias amplification is very serious, and one of the biggest reasons why the effects of these biases are so hard to detect and why the proponents of extreme reductionistic positions can be so resistant to recognizing potential counterexamples to their position. Whatever can be said for theories or paradigms as self-confirming entities (and much that has been said is too excessive, and would render progressive science impossible), as much and perhaps more can be said similarly for heuristics. Indeed, I suspect that most of the blame and criticism of theories in this regard is more accurately laid at the doorstep of the heuristics used by those applying these theories and extending these paradigms.

Consider how this could work. Heuristics 1 and 4, applied in the early stage of an investigation, give apparent conceptual and theoretical reasons for locating a phenomenon of interest (say, that an organism has a given fitness in a given environment) as having causes primarily or wholly within the system under study. In the process of model building, the environment may be simply described (*e.g.*, as totally constant in space and in time, a frequent assumption in population genetics) that relevant variables (and the possibility of their variation) are ignored (Bias 5), further leading to and being reinforced by tendencies not to observe variation in relevant environmental variables (Bias 7) and to make efforts to assure (or, too often, merely to assume) that they are constant as controls in the experimental analysis of the system (Bias 8). Any failures in the model are then assumed to be caused by a failure to model intrasystemic interactions in sufficient detail (Bias 6), leading to another cycle, beginning with biases 1 and 4 applied to the properties and phenomena which were anomalous for the first model. This may result in further simplifications in the environment to offset the loss in analytical tractability arising from the increased internal complexity now assumed, or it may result in focussing in on a particular subsystem to be modelled in further detail, with much of the rest of the system now becoming part of the

systematically simplified, ignored, and controlled environment. If even a part of this scenario or one like it is correct, we should not be surprised if quite remarkable failures went undetected for appreciable lengths of time. At present this remains just a hypothetical scenario, probably only one of many possible scenarios for producing this result. It would be very difficult to establish that the whole scenario, or one like it, was played out in any given case, in part because of the practice of not describing chains of hypothetical reasoning or discovery in scientific papers. Moreover, the practitioners are usually not themselves aware of the microstructure and background presuppositions of their reasoning processes, a fact which has bedevilled attempts to use protocols in which experimental subjects try to describe their reasoning processes as a basis for constructing theories of problem-solving behavior even for much simpler tasks (see Newell and Simon, 1972). Nonetheless, the scientific literature does contain suggestive evidence of several of these heuristics in operation, and it could be hoped that future research would turn up more. A remarkable example of cumulative and systematic biases was unearthed by the work of Michael Wade on the models of group selection, which will be discussed after some preliminary discussion of the notion of a unit of selection in the next section.

6. DARWIN'S PRINCIPLES AND THE DEFINITION OF A UNIT OF SELECTION

Charles Darwin's argument in *The Origin of Species* is adumbrated[5] by R. C. Lewontin (1970, p. 1) as a scheme involving three essential principles:

1. Different individuals in a population have different morphologies, physiologies, and behaviors (*phenotypic variation*).
2. Different phentoypes have different rates of survival and reproduction in different environments (*differential fitness*).
3. There is a correlation between parents and offspring in the contribution of each to future generations (*fitness is heritable*).

Where (and while)[6] these three principles hold, evolutionary change will occur. Lewontin argues not only that these requirements are necessary for evolution to occur, but also that they are sufficient. They also embody that is generally regarded as Darwin's major contribution over prior evolutionists in that they specify a mechanism, natural selection, which produces this change.

Mechanism or not, these principles specify very little about the units which must meet these conditions. Although they are specified in terms of

phenotypes and their properties (a form appropriate to Darwin's original theory, and one to which modern evolutionists still pay lip service), Lewontin immediately applies them to genes (the units of the neo-Darwinian theory, under the impetus of Weismannism). Lewontin exploits the fact that these requirements say little about the units which must meet them, to argue that selection can operate – simultaneously and in different directions – on a variety of units (the unspecified individuals) at a number of levels of organization. In his view, he discusses selection processes at the micro- and macromolecular levels, and as operating on cell organelles, cells (in the immune system, in developmental processes, and, he could have added, in cancer), gametes, individual organisms, varieties of kin groups, populations, species, and even ecological communities.

These principles give necessary conditions for an entity to act as a unit of selection, as well as necessary and sufficient conditions for evolution to occur. The three conditions must all be met by the same entity, in a way that can be summarized by saying that entities of that kind must show *heritable variance in fitness*.[7]

These conditions fail to be sufficient for the entity to be a unit of selection, however, for they guarantee only that the entity in question is either a unit of selection *or is composed of units of selection*. A further condition, which is sufficient, is given in the following definition:

> A *unit of selection* is any entity for which there is heritable *context-independent* variance in fitness among entities at that level which does not appear as heritable context-independent variance in fitness (and thus, for which the variance in fitness is *context-dependent*) at any lower level of organization.

Much of population genetic theory involves the notion of additive variance in fitness. It is this quantity which, in Fisher's fundamental theorem of natural selection (Fisher, 1930) determines the rate of evolution. To say that variance in fitness is totally additive is to say that the fitness increase in a genotype is a linear function of the number of genes of a given type present in it. But this entails that the contribution to fitness of a given gene whose effect on fitness is totally additive is independent of the genetic background in which it occurs, which is to say that the variance in fitness is context-independent. Additivity is thus a special case of context-independence. It is assumed for reasons of analytical tractability, but the properties which flow from this assumption derive from its relation to context-independence.

One very important result follows when this assumption holds at a given level of organization. *If variance in fitness is totally additive at a given level of organization over a given range of conditions on the environment and the system, then, under those conditions there are no higher-level units of selection*! This is true because fitness of any higher level unit is then a totally aggregative or mass effect of the fitnesses of the individual entities at that level of organization. With no context-dependence of fitness, the *organization* of these units into higher-level units does not matter. There are no epistatic interactions to tie complexes of these entities together as units of selection. The higher level unit is totally reducible in its effects to the action of various lower level units, acting in a context-independent manner.[8]

It may be that this assumption (a product of bias 4 or 5 applied at the level of the gene to increase the analytical tractability of the model) is one of the major reasons contributing to the plausibility of Williams's reductionistic vision. It is clear that once this assumption is made, it becomes plausible to attribute adaptation (and thus fitness) "to the effect of selection acting independently at each locus" (Williams, 1966, p. 57) and leads naturally to regarding fitness as a property of genes (a case of bias 1). It is also true that many or most population geneticists believe and argue (as James Crow has, in personal conversation) that most variance in fitness is additive – presumably, at the level of the contributions of individual genes. This is an empirical claim and represents a view not shared by all population geneticists. Sewell Wright has systematically argued throughout his professional life and his magisterial four-volume treatise that the opposite is true, that epistatic interactions are all pervasive and important (personal conversation; see *e.g.*, Wright, 1968, Chapter 5, especially pp. 71–105). Michael Wade's current research indicates the importance of epistatic interactions at the *individual* level (that is, between individuals in populations) in group selection (personal conversation; see Wade and McCauley, 1979, and McCauley and Wade, 1979). What is clearly true is that biases 7, 8, and 9 would in general contribute substantially to failures to detect nonadditive variance if it exists because of artificially induced constancies in or ignorance of environmental conditions capable of producing nonadditive components of variance in fitness.

To summarize then, if variance in fitness at a given level is totally additive, the entities of that level are composed of units of selection, and there are no higher level units of selection. If the additive variance in fitness at that level is totally analyzable as additive variance in fitness at lower levels, then the entities at that level are composed of units of selection at these lower levels, rather than being units of selection themselves. To put it in terms of Salmon's (1970) analysis of statistical explanation, the higher level units of selection as causal factors are then 'screened off' by the lower level units of

selection. In their causal effects, they are then 'nothing more than' collections of the lower level entities, and any independent causal efficacy is illusory. This is a necessary and sufficient condition for the truth of Williams's genetic reductionism.

But in general, we would expect this partitioning of variance in fitness into additive and nonadditive components at different levels to show a number of levels – genes, gene complexes, chromosomes, individuals, even groups – at which additive variance at that level appears only as nonadditive variance at lower levels. There are units of selection at each level at which this occurs, and if it does, genetic reductionism and determinism are false.

7. WADE'S REVIEW OF THE MODELS OF GROUP SELECTION

For groups to act as units of selection, they must show heritable context-independent (or additive) variance in fitness[9] which is not merely a summative redescription of additive components of variance in fitness to be found at lower levels of organization. Groups must meet Darwin's three principles as well as the additional constraint of a context-independent component of fitness to so qualify.

Organisms meet Darwin's principles by having phenotypic differences (variability) which are heritable, and which have a differential effect on their survival and/or reproduction. Similarly, there may be differences of group structure and interaction (variability of group phenotype) which are transmitted to offspring groups or migrant propagules (group heritability), and which affect the rate of survival and/or reproduction of groups. The heritability of group variability may be either genetically or phenotypically transmitted. The latter results in models for cultural transmission and evolution. The former results in models for the genetic transmission of group traits and biological evolutionary models of group selection. Wade's results concern these models, in which what is inherited, to a greater or lesser degree, is the set of gene frequencies of a group's gene pool. Thus, if migration from a group occurs at random with respect to the genotypes of the individual migrants, the fact that the migrants are drawn from a given group will confer a kind of heritability or correlation between the gene frequencies of the parent population and the gene frequencies in the migrant propagule. (Indeed, this may be true to some extent even if migration from the group is not at random with respect to genotype.) This, together with a differential rate of production of migrant propagules by groups of different genetic compositions or differential rates of survival of such groups or both should make group selection a reality.

The various mathematical models of group selection surveyed by Wade all admit of the possibility of group selection. But almost all of them predict that group selection should be a significant evolutionary factor only very rarely, because they predict that group selection should have significant effects only under very special circumstances – for extreme values of parameters of the models which should be found in nature only rarely. Wade undertook an experimental test of the relative efficacy of individual and group selection – acting in concert or in opposition in laboratory populations of the flour beetle, *Tribolium*. This work (reported in Wade, 1976 and 1977) produced surprising results: group selection appeared to be a significant force in these experiments, one capable of overwhelming individual selection in the opposite direction for a wide range of parameter values, apparently contradicting the results of all of the then extant mathematical models of group selection. This led Wade to a closer analysis of these models, with results reported in Wade (1978), and described here.

All of the models surveyed made simplifying assumptions, most of them different. Five assumptions however were widely held in common: of the twelve models surveyed, each made at least three of these assumptions, and five of the models made all five assumptions.[10] Crucially, for present purposes, the five assumptions are biologically unrealistic and incorrect, and each independently has a strong negative effect on the possibility or efficacy of group selection. It is important to note that these models were advanced by a variety of different biologists, some sympathetic to and some skeptical of group selection as a significant evolutionary force. *Why then did all of them make assumptions strongly inimical to it*? Such a 'coincidence', highly improbable at best, cries out for explanation. I will attempt to offer some explanations after I have presented and discussed some of the assumptions. These assumptions are given in Wade (1978, p. 103):

(1) It is assumed that the frequency of a single allele within a population can produce a significant change in the probability of survival of that population, or in the genetic contribution which the population makes to the next generation.

(2) All populations contribute migrants to a common pool, called the "migrant pool" (Levins, 1970), from which colonists are drawn at random to fill vacant habitats.

(3) The number of migrants contributed to the migrant pool by a population is often assumed to be independent of the size of the population. Thus, the frequency of an allele in the migrant pool

can be represented by the mean allele frequency of all contributing populations.

(4) It is assumed, often implicitly, that the variance between populations (which is a prerequisite for the operation of group selection) is created primarily by genetic drift within the populations and, to a lesser extent, by sampling from the migrant pool.

(5) Group and individual selection are assumed to be operating in opposite directions with respect to the allele in question. In short, the allele is favored by selection between groups but disfavored by selection within groups.

The first assumption replicates Williams's reductionistic approach by assuming that the trait under selection at the group and individual levels is a single-locus trait, rather than one with a polygenic basis. However, Wade's experiments rule out this possibility for the trait under selection, population size. In these experiments, a number of replicate populations given the same selection treatment (selection for increasing size at both group and individual levels; selection for decreasing size at both levels; or selection for increasing size at one level and decreasing size at the other) were assayed to determine a number of demographic parameters of the population. These demographic parameters included fecundity, fertility, body size, developmental time, and cannibalism rates for various developmental stages on other developmental stages, all factors having a known and modelable effect on population size, and on the number of different organisms of different age classes as a function of time. (The demographic effects of changes in these variables are visualized in the comparison of the number of individuals of different life stages in high productivity and low productivity populations as a function of time in Figure 4 of McCauley and Wade, 1979.) The results of this assay showed that replicate populations drawn from the same genetic stock, and given the same selection treatment, to which they responded in the same way at the level of the macroscopic trait, population size, achieved this response by different combinations of changes in the underlying demographic parameters affecting population size![11] This means that even if each of these underlying demographic parameters is affected by a single-locus trait (which seems unlikely), *population size cannot be a single locus trait because similar values of it are produced by simultaneous independent variations in the underlying demographic variables*. Thus none of the single-locus models (nine out of the twelve) can adequately describe the selection processes in this experiment. In virtue of the increased dynamical complexity and richness of two-locus

systems as compared with one-locus systems (see above, Section 4, and Roughgarden, 1979, Chapter 5) and the further poorly understood complexities of systems involving more than two loci (see Lewontin, 1974, Chapter 6 and Maynard Smith, 1978, Chapter 5), this is a very significant failure. Little if anything learned from the single-locus case is generalizable to cases involving two or more loci.

The second assumption, the analytical device of a 'migrant pool', is, if anything, more serious. In this assumption, all migrants contributed by any population are thrown into a common pool (see Figure 3) from which all new

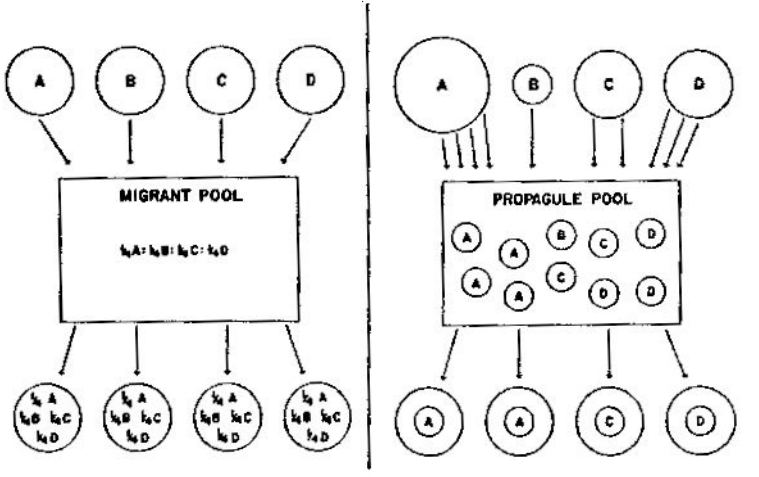

Fig. 3. Diagram illustrating the differences between a migrant pool and a propagule pool. (Reprinted from Wade, 1978, Figure 1, p. 109, with permission of the *Quarterly Review of Biology*.)

founding populations are drawn. This assumption may not be unrealistic in some cases in the five intrademic selection models (it is equivalent to assuming panmixia or random mating within populations), but is surely radically unrealistic for any of the seven traditional models, where it is equivalent to assuming panmixia throughout the entire range of a species. If there is any systematic genetic differentiation throughout the range of a species, and if the averge one-generation migration distance of an individual of that species is appreciably less than the dimensions of its range, then there will be mating which is assortative (or nonrandom) merely because of the limited distance a migrant can travel, thus falsifying this assumption. In general, parent

populations will contribute migrants to founding populations in their immediate vicinity, and this assumption will be seriously incorrect.

The seriousness of this simplifying assumption can be better seen by exploring its analogies with the theory of blending inheritance at the individual level. In 1868, Darwin proposed his 'provisional hypothesis of pangenesis' in which large numbers of gemmules secreted by the various cells of an organism were combined in sexual reproduction in such a way that the characters of the offspring (produced by equal contributions of gemmules from both parents) were an intermediate blend of the characters of the parents. This theory drew rapid and searching criticism (as it appeared in the fifth edition of Darwin's *Origin of Species*) by Fleeming Jenkin (1867), who pointed out that with such a 'blending' mechanism of inheritance, the variation in a population would be rapidly attenuated until the population was essentially homogeneous. With no variation for selection to act upon, evolution would rapidly come to a halt.

With the rise of Mendelism (Mendelian segregation prevents significant blending in the 1-locus case, and limits its effects in the multi-locus case[12]) blending theories were rapidly forgotten, and almost without exception (one being Wallace's excellent discussion in Wallace, 1968, pp. 61–65), their characteristics and consequences are forgotten, ignored, and misunderstood.[13]

To R. A. Fisher, however, the avoidance of blending inheritance and its consequences for the loss of genetic variance was a *sine qua non* for the possibility of evolution. He began his ground-breaking treatise (Fisher, 1930, pp. 1–4) with a discussion of the character and consequences of a blending mode of inheritance. To Fisher, Mendel had clearly made the world safe for Darwinism, and he turns somersaults in a marvelously Whiggish rewriting of history to try to show that Darwin didn't really take blending inheritance too seriously.

Whatever the merit of that attempt, he makes the consequences of blending inheritance very clear. Fisher's fundamental theorem of natural selection says that a measure of the rate at which gene frequencies change is directly proportional to the additive genetic variance in fitness. No additive genetic variance in fitness means no gene frequency change, and no evolution. In those beginning pages, he derives (or actually, claims the derivability of) a formula for the rate of loss of variance resulting from mating under blending inheritance. In each generation, the variance is attenuated by a factor of $\frac{1}{2}(1 + r)$, where r is the correlation between parental genotypes. Fisher expected this correlation to be small (r ranges between -1 and $+1$, and is 0 in a randomly mating population). Thus, he argued that the existing genetic variance in a population would be approximately halved in successive generations. With no production of new variation, evolution would go as far in a given generation as it would in all successive generations, in accordance with the series $1 + \frac{1}{2} + \frac{1}{4} + \ldots$ (see Wallace, 1968).

Blending inheritance at the individual level is analogous to the assumption of the panmixia of migrants, or of a mating pool at the group level. Wade and I (Wade and Wimsatt, in preparation) have rederived and extended Fisher's results in a manner applicable to the case of group selection. Blending inheritance at the individual level involves two parents, each making equal contributions. The group case may involve contributions of migrants from n parental populations, in possibly unequal contributions, $w_1, w_2, \ldots, w_n$, which sum to 1. In a deterministic model which neglects sampling error in drawing migrants from parental populations, and in which the draw does not affect the gene frequencies in the parental population (requiring, in effect, infinite population size of both parental and migrant propagule populations with the migrants constituting a negligible proportion of the parent populations), a close analogy with Fisher's result can be derived. In this case, if the correlation between any two parental populations is constant at r, then the factor by which variance is attenuated in successive generations, α, is given by equation 1:

$$(1) \qquad \alpha = r + (1 - r) \Sigma w_i^2$$

This reduces to Fisher's derived factor of $\frac{1}{2}(1 + r)$ if there are two parents making equal contributions. With n parents making equal contributions, and no correlation between them, ($r = 0$), $\alpha = 1/n$. This is a 'worst case' assumption, but it is exactly what is assumed in the 'migrant pool' models! The assumption of the migrant pool with any significant number of populations guarantees that the variance goes to 0 essentially immediately.

Suppose that variance is being created anew in each generation by some mechanism or mechanisms, at a rate V_0. Then the equilibrium pool of variance (which will determine the rate of evolution) is that amount of variance for which variance is being lost at the same rate at which it is being created, which happens when:

$$(2) \qquad V = [1/(1 - \alpha)]\ V_0$$

Under the migrant pool assumption, this is simply $[n/(n - 1)]\ V_0$, which is essentially V_0 for n large. In effect, there is no pool of variance, but selection can only act on the variance created anew in each generation.

It can be seen by inspection of equations (1) and (2) what conditions will allow the accumulation of an appreciable amount of variance. This happens when α is close to 1, which is true if either there is high correlation between parent populations (r close to 1), or if one population predominates in contributing to a given migrant propagule (Σw_i^2 will be close to 1 if and only if w_i is close to 1 for some i). In the migrant propagule model of Wade (1978), the latter condition holds, since all of a migrant population comes from a single parent population, and $\alpha = 1$ (in this deterministic model, which doesn't allow for sampling error, and thus does not exactly apply to Wade's case).[14] In more realistic models for the diffusion of migrants in a cline, the former condition would tend to hold. So Wade's models, and realistic population models which don't make the migrant pool assumption, would be expected to allow much more rapid evolution at the group level. Plausible values of the parameters r (or r_{ij}'s in a more general treatment) and the w_i's suggest rates of evolutionary change at the group level which are easily 10 to 100 times greater than expected under comparable circumstances with the migrant pool assumption.[15]

Why the seriousness of the migrant pool assumption should have been overlooked, and the significance of how it was finally detected and analyzed, relate to the use of heuristics, and will be discussed in the closing section.

The third assumption, that parent populations contribute the same number of migrants to the migrant pool, independent of their size, in each generation, was also made for reasons of analytical tractability given by Wade (1978). Its effects however are far more immediately obvious, and thus it is similarly more anomalous that the assumption would ever have been made. In models where it is made, differences in reproductive output of different populations is due entirely to the populations' surviving for a different number of generations, since all alike have the same reproductive output in generations in which they are surviving.

Such an assumption would never have been made in models of selection at the individual level. Darwin was acutely aware of the importance of differential reproduction as a selective force, one overshadowing in its potential intensity the effects of differential survivorship, and other evolutionists since have generally retained this awareness. It is doubly mysterious as an assumption at the group level, since presumably one of the primary ways in which increased group fitness could be manifested would be through increased production of individuals, leading to increased population size and increased migration of individuals once maximum population size is attained. One would thus expect a strong correlation between population size and migration

rate, rather than the constancy assumed in these models. The association between size and reproductive rate is even more direct than it is for most cases of selection at the individual level. This assumption also has an enormous effect under some conditions, particularly those in which the average population survives for several or more (population) generations in which the averge output of migrant propagules is appreciably more than one per generation. This can be seen by looking at the coefficients in Table 2, which gives the ratio in per generation reproductive rates required of two replicators, one

TABLE 2

Ratio of intrinsic growth rates, r, of shorter-lived replicator (r_b) to longer-lived replicator (r_a) that shorter-lived replicator needs to offset shorter lifetime.**

lifetime of replicator (generation)	Intrinsic growth rate (per generation) of a replicator with a lifetime of 1000 generations					
	1.001	1.01	1.1	2	10	100
1000	1	1	1	1	1	1
100	1.0062	1.0032	*	*	*	*
10	1.071	1.067	1.033	1.001	*	*
5	1.148	1.143	1.101	1.0062	*	*
4	1.188	1.183	1.139	1.015	*	*
3	1.259	1.254	1.205	1.040	1.0003	*
2	1.413	1.403	1.352	1.118	1.005	*
1	1.998	1.990	1.909	1.500	1.100	1.010
.9	2.158	2.149	2.062	1.611	1.141	1.017
.75	2.518	2.507	2.405	1.863	1.244	1.042
.50	3.997	3.980	3.816	2.914	1.733	1.210
.25	15.988	15.921	15.260	11.485	5.958	3.001
.10	1023.25	1018.92	976.46	728.44	345.95	133.18
.01	1.27×10^{30}	1.26×10^{30}	1.21×10^{30}	8.97×10^{29}	4.04×10^{29}	1.30×10^{29}

* Added increase is less than 1 in 10^4, Fisher's rough lower limit for selective differences to be significant.

** selection coefficient is $r-1$.

r short (r_b) calculated from r_{1000} (r_a) from the relationship $r_b = \{[(r_a)^a - 1]^{b/a} + 1\}^{1/b}$ and from the approximation $r_b = [(r_a)^b + 1]^{a/b}$ when $(r_a)^a$ is greater than 1×10^{100}.

with very long lifetime (1000 generations) and with different assumed reproductive rates given across the horizontal dimension, the shorter with a lifetime given in the vertical column and with a per generation replication rate

such that the shorter lived replicator will have the same long-range replication rate as the longer lived one. Thus the shorter lived replicator must increase its per generation reproductive rate above that of the longer lived replicator (indicated by the amount that the ratio of reproductive rates exceeds 1) in order to offset the effect of its shorter lifetime. Comparisons in which the longer lived replicator is not so long lived (say ten generations) can be gotten by dividing the appropriate coefficient of the shorter lived replicator (say 1.0062 for comparison with a five-generation replicator if the ten-generation replicator is having two offspring each generation) by its appropriate coefficient (1.0001, in this case), to get the ratio (1.0061, in this case) of per generation reproductive rates for the five- and ten-generation replicators.

What should be obvious from looking at this table is that unless one is comparing replicators which survive for a relatively short time *and* have very low reproductive rates, almost negligible differences in reproductive rate more than offset even quite substantial differences in lifetime. This is so because in this range of parameter values, changes in lifetime have a roughly linear effect on the outcome, whereas changes in reproductive rate have a roughly exponential effect. Thus a small increase in reproductive rate offsets a large decrease in lifetime.[16] Given this fact, it is a bizarre and dangerous simplification in a model to assume that per generation reproductive rate (a variable to which the net reproductive rate, which determines the intensity of selection, is very sensitive) is constant across populations, and that differences in selection are due entirely to differences in the lifetimes of different populations (a variable to which it is quite insensitive). Presumably, group selection would often, even usually, act to optimize both the reproductive rate and the survival time of populations, but these considerations, suggest that the former variable bears watching much more closely than the latter.

I will make no lengthy comments on the last two simplifications discovered and discussed by Wade. Clearly the fourth rules out spatial heterogeneity as a significant selective force at either the individual or group levels, an assumption which also exerts a bias against higher-level units of selection which will be discussed on another occasion.

The fifth simplification, the assumption that group and individual selection are operating to change gene frequencies in opposite directions, has two interesting features.

The first is that it almost certainly has its origin not in any arguments about what would be true in nature, but in the joint action of a consideration of testability and a simplifying assumption. The consideration of testability is that, because of the complexity of interaction of fitness components and the

difficulty of determining the relevant parameter values (see Lewontin, 1974), it would be helpful in determining the efficacy of group selection if we could find a trait whose presence clearly signalled the operation of group selection, because it would be selected *against* at the individual level and thus could *only* owe its presence to group selection. This does *not* mean that group selection would usually or generally tend to be opposed to individual selection in nature. Nonetheless, it was probably responsible for the concentration of analytical models on circumstances designed to investigate this condition. In the context of models of two alleles at one locus, the natural way to implement this condition is to assume that the effect of selection at one level was to increase a given gene frequency and that of selection at the other level was to decrease it. In the context of such simplified models, the move from traits to genes or genotypes is easy — all too easy as recent sociobiology has shown — but this way is fraught with error, as a longer discussion on another occasion (see Note 7) will show.

It is worth noting only how implausible the assumption that individual and group selection are opposed becomes once multi-locus models are considered. The effects of selection may generally be described as a vector in which each component is the change in one of the stable variables (*e.g.*, frequencies of genes, gametic, or zygotic genotypes) describing the population. *Only in a phase space of one dimension, such as that characterizing the model of two alleles at one locus, are change vectors constrained to life in the same or in opposite directions. In spaces of any higher dimensionality, the probability that they will be identically or oppositely directed is of measure zero, and the resultant of the two vectors may similarly lie in any direction whatsoever.* Any residual plausibility of this assumption is clearly an artifact of being guided only by the simplest possible model of evolutionary change.

But then if group selection no longer has to overcome forces of individual selection to which it must be opposed, it matters little what or how strong are the selective forces acting at the individual level in evaluating the possibility that group selection can be efficacious. This is particularly detrimental to many of Williams's (1966) arguments. In a multi-dimensional space, even relatively weak selective forces at the group level, when added to relatively strong forces at the individual level, can change the resultant selection vector sufficiently to cause evolution towards an alternative adaptive peak than that which might be achieved by individual selection acting alone. The richer dynamics and greater dimensions of a multi-dimensional phase space produces the possibility of a wide variety of interactions among selection forces at different levels. These interactions are surely poorly understood at this time.

But we cannot hope to understand them if we don't even detect them. We fail to do so, it appears, because of biases which are almost perceptual in character.

8. HEURISTICS AND THEIR BIASES: SOME AMELIORATIVE REMARKS

The simplifications Wade discusses were almost all (with the possible exception of the fifth) made to improve analytical tractability. Why their biasing effects should not have been noticed is a difficult and probably multifaceted problem. I will here discuss some of the considerations which seem to me to be most salient:

(1) *Inertia.* Some assumptions have a time-honored status, in that they have been made by almost all past models. As a result, unless the model or phenomenon in question itself seems to immediately point to the need for relaxing one of these assumptions, it will be taken for granted, especially since each of these assumptions almost invariably involves advantages of increased analytical tractability. This may relate to the anchoring bias of Tversky and Kahneman (1974), but it seems equally likely that 'anchoring' and 'inertia' are broad phenomenological categories that cover a multitude of sins committed for a variety of reasons. Particularly well entrenched is the assumption of panmixia (equivalent to the mating pool assumption, and *sufficient to guarantee that the mode of inheritance at any higher level of organization is a form of extreme blending*). An assumption that is well entrenched in the theoretical structure is one that is widely used in and at least apparently essential to the derivation of many other results. These dependency relations can lead to a reluctance to give up an assumption even if it is widely known to be problematic, or even if it is generally believed to be false. The assumption of the transitivity of preferences in decision theory, or of the transitivity of fitness in evolutionary theory, both known to be false when choice or fitness is a function of a number of variables, seems to me to be such a case. The assumption of panmixia is almost certainly another. A third, similarly entrenched assumption in quantitative genetics is the assumption that variance in fitness is totally additive. Both this and the assumption of panmixia are absolutely inimical to the existence or significance of higher-level units of selection.

(2) *Perceptual focus.* Given the centrality of reproductive rate in virtually all evolutionary models and in the structure of evolutionary theory, Wade's third simplifying assumption seems almost unintelligible in almost any

circumstances. The only way that I can think of to rationalize it is as follows. model-building activity is performed against a background of presumed mechanisms operating in the interaction of presumed units. If the presumed units are very well entrenched in a given area, there is a strong tendency to describe and to think about even phenomena at other levels of organization in terms of these units. In traditional evolutionary theory, and even at present, the most obvious unit is the individual organism – the unit which our everyday thought and our perceptual apparatus naturally predisposes us to consider. Most of our everyday interactions (as well as those of most other organisms) that call for voluntary action are pairwise interactions with one other organism at a time, so this bias is evolutionarily well founded. Consequently, there is a strong tendency to see and to talk about groups of organisms as *collections of individuals, rather than as unitary entities.* This is true even for colonies of social insects, whose interdependencies extend even to reproductive specialization, making the metaphor of the colony as an organism perhaps more revealing in evolutionary terms than the view of it as a collection of organisms. Hull (1978) has found similar biases in his arguments that species must often be conceived of as individuals in evolutionary contexts. A quick review of Williams (1966) reveals that even in the context of discussions of group selection, groups are usually described as collections of individuals, and it is my impression that this tendency is widespread throughout the literature. But, *in the context of the group selection controversy, description of an assemblage of units as a collection is a theory-laden description, since it suggests that it is an aggregate. The only time it is appropriate to describe it as an aggregate is when the fitnesses of its components are context-independent* (see Section 6) *or are additive. But as we saw above, this is a sufficient condition for their not constituting a higher-level unit of selection*! Perceptually, the focus on individual organisms prevents us from at the same time seeing the groups as individuals. If we do not see the groups as individuals, then we do not see that assuming that each group contributes an equal number of migrants is equivalent to assuming that there is no variance in reproductive rate. The bias, I suspect, is a perceptual one.

Similar remarks apply to seeing panmixia as equivalent to a form of blending inheritance at the group level. Blending inheritance is traditionally viewed as applying only at the individual level. But once the group is seen as an individual, a view that emerges particularly strongly in Wade's work, even in his earliest paper (Wade, 1976, 1977), the analogy is immediate. I saw the analogy when I read an early draft of Wade (1978) at the same time as I was teaching Darwin's blending theory and its criticisms. The phenomenon was

reproducible: three months later, Ross Kiester was teaching Darwin and independently pointed out the analogy to Wade. (Seeing the analogy also requires, I suppose, more than a passing acquaintance with blending inheritance.)

The recent focus on the gene (rather than the individual) as the unit of selection has introduced its share of perceptual problems and biases. Williams (1966) and Dawkins (1976) are full of 'perceptual shifts' back and forth from the genetic to the individual level, but both try to maintain the primacy of the former with consequent biases in their description of the complexities of gene interaction.

Wade (1979) has shown a marvelous case of perceptual bias in the foundational work of Hamilton on kin selection theory. In this theory, selection is seen as maximizing an individual's 'inclusive fitness', in which contributions to the fitness of relatives who share genes are included, weighted by their degree of relationship. Hamilton is reported to have described this as "the gene's eye view of evolution," and in any case the description fits. Hamilton's theory involves the assignment of selection coefficients to individual genes, and makes no reference to the genotypes in which they occur, or to which genotypes mate with which to produce offspring. Wade (1979) advances one- and two-locus models in which this minimum amount of detail is added, and in which he is able to prove an analogue of Fisher's fundamental theorem of natural selection, according to which the rate of evolutionary change is proportional to the *between family variance in fitness*. This result shows that *in a more realistic model of kin selection, the unit of selection is not the individual gene, but the mating pair, Wade's 'atomic family'*. This result could not have been derived or even seen in Hamilton's model, because the simplifications of the environment of the gene were such that Hamilton had no structure in his model on which to hang the upper-level phenomena. Mating pairs do not occur in his model, and there is no way of putting them into it. This is a paradigmatic example of Bias 5 in Section 5 above, but also shows the perceptual or quasi-perceptual character of the way in which this bias operates.

(3) *Perceptual reinforcement*. A factor discussed above is the reinforcement of different biases or heuristics. When a force must be greater than a given magnitude to have an effect, several biases tending to underestimate this force may have a cumulative effect that none of them could have alone. But biases can be mutually supportive in another way also suggested there. One bias may act in such a way as to hide the fact that another bias *is* a bias, and conversely. Detailed documentation of these cooperative effects of biasing assumptions in population genetics remains largely a task for future analysis,

but a promising one. The assumptions of panmixia and additivity of fitness variance are plausible candidates for inspection, both because of their ubiquity and also because each carries with it a relatively easily intuited picture of the unstructured environment of the unit of analysis. The assumption that fitnesses are constant in space and in time, and independent of the density or of the relative frequency of the relevant units (usually assumed to be genes) also demand closer inspection, as each is a source of potential systematic bias against recognizing higher level units. Near decomposability assumptions are important tools which are easily misused, here as elsewhere. Some further biases which may be relevant in this context are discussed in Wimsatt (1980), but most of the work is yet undone.

How, aside from analyzing specific cases of bias, and identifying biases more generally can something be done to correct for their effects? There are at least two plausible candidate procedures:

(1) The first is what Levins (1966) has called the search for robust theorems. To counteract the biases of any given model, he suggests building families of alternative models of a given phenomenon which differ in their simplifying assumptions. The models will vary in their consequences and predictions, reflecting the variety of their assumptions, but there may be consequences which are true across all of the models. These Levins calls 'robust theorems', results which are independent of the details of any particular model. Thus, says Levins, "Our truth is the intersection of independent lies" (1966, p. 126; see also his 1968, pp. 6–8, for further elaborations of this approach).

This method has two disadvantages, neither fatal, but both of which are worthy of note. The first is that we are often in situations in which we do not have even one model, much less two or more, to compare their results. Population genetics and ecology has a richness of mathematical models, unlike some other disciplines, but even here, many areas do not have models with sufficient overlap to be able to compare the results. A further contributing problem arises when the assumptions made in deriving the models are not explicit, since it is then not possible to tell how much one can trust the robustness of a result, basically for the second reason.

A second caution is necessary because it is not always possible to tell when the models are *in fact* independent. If a number of models each make a possibly well-disguised but in any case unnoticed assumption, any theorems in common may simply be a relatively direct consequence of the shared assumption or assumptions. In this case, the results will generally be assumed to be quite robust – illegitimately as it turns out. Indeed, just this seems to

have happened in the group selection controversy. Until Wade's review, where the assumptions the various models had in common and their consequences were made explicit, the one result which seemed to be robust was that group selection could be efficacious only rarely and under very special circumstances, and probably the majority of evolutionary biologists still believe this to be the case.

Another approach, generally consistent with robustness analysis (see Levins, 1968, pp. 6–8) and suggested by the heuristic use of means-end analysis and identifications in inter-level reductive explanation (Wimsatt 1976*a*, pp. 231–237, 1976*b*, Section 8) could be called 'multi-level reductive analysis'. Even if perceptual focussing may leave one blind to the biases of the nine heuristics of Section 5 as applied at any one level, these same heuristics will have biases leading to *different* simplifications if they are applied to the same system at a different level of organization, *simply because the system-environment boundary has changed*. Simplified models of group selection may thus suggest particular structural features of the environment of the individual which should be included in models of individual selection, just as, *e.g.*, an analysis of the structure of the genotype and/or mating system of the individual may suggest important internal constraints on models at the group level. Simultaneous multi-level modelling may thus eliminate the biases of proceeding at only one level, but a final caution is required: this is likely to work only to the extent that the phenomena and entities of a given level are taken seriously in their own right. Seeing them merely as an extension of another level, be it lower or higher, will merely preserve the perceptual focus of that other level, and most biases will go undetected. This is merely another expression of a view I have argued for before (Wimsatt 1976*a*, 1976*b*, 1978). Now for pragmatic as well as for theoretical reasons, reduction in science is better seen as the attempt to understand the explanatory relations between different levels of phenomena, each of which is taken seriously in its own right, than as an unending search of firm foundations at deeper and deeper levels in which, as Roger Sperry so aptly put it (quoted in Wimsatt, 1976*a*), " ... eventually everything is explained in terms of essentially nothing."

Department of Philosophy,
Committee on Evolutionary Biology, and
Committee on Conceptual Foundations of Science
The University of Chicago

ACKNOWLEDGMENT

This work owes a great deal more to many people than is reflected in the footnotes: To Richard Lewontin for his leading me past the most elemental understandings (and sometimes misunderstandings) I brought to my study of population genetics as a post-doctoral fellow in his lab, and for his early advocacy of the importance of gene interactions and higher level units of selection in evolution, which, as with many other things he has taught me, had a formative effect on my present views. To Richard Levins I owe a similar debt, for his tutelage in mathematical ecology and the analysis of complex systems, for his friendly tolerance of my early naive and principled reductionism, while he gradually taught me elements of a richer vision. To George Williams, whose book was an education to me and to many others, formulating for the first time a clear reductionistic vision and many arguments supporting that view, which I continue to respect, even though I now believe it to be mistaken. More recently, I have learned a great deal from discussions with Ross Kiester and Mike Wade about model building and experimental design, as well as a lot of fascinating and useful biology, and Wade in particular has done a lot to shape my views on group selection. I have profited also in particular from discussions with David Hull, but also with Bob McCauley, Bill Bechtel, Bob Richardson, Elliot Sober, and Mary Jane West-Eberhardt, the participants of a Midwest Faculty Seminar on Evolution at Chicago, and seminar or conference participants at the Leonard Conference at Reno, Denison University, Ohio State, the University of Wisconsin at Madison, and the University of Colorado at Boulder, where versions of this work were presented in 1978–79.

I would like to thank Columbia University Press for permission to reprint Figure 23 and 24 from Lewontin (1974), pp. 274 and 280, and to adapt Table 56 from p. 283; and the *Quarterly Review of Biology* for permission to reprint Figure 1, p. 109, and to quote from p. 103 of Wade (1978). Finally, all of this would have occurred substantially later or perhaps not at all without the support of the National Science Foundation (Grant SOC78-07310 -research).

NOTES

[1] Actually the term 'heuristic' was used earlier and perhaps introduced by G. Polya in his book *How to Solve It* in 1945. The idea of a heuristic procedure has however developed substantially further, and has become one of the central theoretical concepts of artificial intelligence. See, *e.g.*, Nilsson (1971) and Winston (1977) for some more recent discussion of heuristic programming.

[2] On the most obvious reading, in which genotypes are the subsystems containing genes, which interact to produce fitness, which is a property of genotypes which affects the multiplication ratios of the genes they contain, selection and mating would be treated as an intersystemic interaction, with segregation and recombination as intrasystemic interactions. Thus the analogy is not quite exact. In this case, with mating assumed to be at random and the genotypic frequencies in linkage equilibrium, the 'long range' behavior of the system involves the interaction of an intrasystemic force (segregation) and an intersystemic one (selection). This particular decomposition into subsystems (at variance with that required for easy analysis of near decomposability) is necessitated

by the particular structure of Mendelian genetics, which, through mating and differential reproduction, inextricably combines inter- and intra-organismic forces.

Other ways of breaking up the system (*e.g.*, into loci as subsystems, which might be suggested by Williams's remarks) produce similar problems: then recombination (a strong force if there is substantial linkage disequilibrium) is intersystemic, rather than intrasystemic, as it 'should' be. Nonetheless, the partitioning of forces into strong and weak, characteristic of near decomposability analysis is found here also, so there remains an important (and probably the most important) ground of analogy.

[3] The biologist's use of tautology here is looser than the philosopher's, and means roughly a relation which has no empirical content because of the way in which it is used. As such it is related to the vernacular use of tautology as in 'covert tautology' rather than to the logician's sense. I will here use the term as the biologist does.

[4] These judgments are themselves an important source of error, associated with heuristics for cutting the world up into entities, using the robustness or overdetermination of boundaries. These heuristics and examples of their application and misapplication are discussed at length in Wimsatt (1980) where they are particularly relevant in understanding the nature and origin of functional localization fallacies.

[5] These make no mention of the geometric rate of natural increase of organisms and the consequent inevitability of competition for resources (Malthus's observation). But this was a subsidiary argument employed by Darwin to establish the second principle – that different types of organisms had different fitnesses. Darwin needed this *a priori* argument because he had no direct observations of the occurrence of natural selection in nature.

[6] Lewontin applies these principles on a genetic micro-evolutionary scale, and points out that for a population in equilibrium of gene frequencies, however temporary, conditions 2, 3, or both are not met (1970, p. 1). And obviously, if there is only a single allele at a given locus in a population (violating condition 1), no change in gene frequency (or micro-evolution) is possible at that locus.

[7] I have analyzed these conditions and their ramifications in much greater detail in a book manuscript now in process and tentatively to be called *Reductionism, Sociobiology, and the Units of Selection*. Further excellent discussions of related issues can be found in Hull (1978), Sober (1979), and, less directly, Cassidy (1978).

[8] Mike Wade felt that this did not emphasize sufficiently strongly that whether an interaction was additive or epistatic is a function of the relation of the system to the environments in which it is studied. He feels that many studies which purport to show that variance in fitness is additive rather than epistatic suffer from looking at a restricted environment (usually in the laboratory) or range of environments, and that investigation of the system in a wide range of environments would show that many or most of the supposedly additive interactions are in fact epistatic.

It is worth pointing out that the term 'epistasis' is traditionally reserved for interactions between genes *within a given genotype*. But the discussion here naturally suggests an extension to interactions between higher level complexes of genes. Thus, when one speaks, as Wade does, of a *group* phenotype, it becomes natural to describe nonadditive interactions between individuals in the group as epistatic.

[9] The notion of additive variance has an implication that speaking of context-independent variance does not. Speaking of additive variance implies the context of a larger unit in which more than one of the smaller units which contribute to fitness will co-occur, so that their contributions will add. Thus genes which show additive variance will occur in genotypes. In the case of two allels at one locus, this condition is met if the fitness of the heterozygote *Aa* is exactly halfway between the fitness of the two homozygotes, *AA* and *aa*. Talking about groups as units of selection may not imply a larger conspecific group whose fitness they contribute to, and in this case (and other similar cases), it is preferable to talk about *context-independent fitnesses* of groups rather than of *additive contributions to fitness* of groups.

[10] Within the 'traditional' models, the record was even worse. Five out of seven of the models made all five of the assumptions. The structure of the intrademic models, a newer development widely heralded as improving the case for group selection, required dropping one of the assumptions.

[11] Thus, population size is a case of a functional or supervenient property, in the sense widely discussed in philosophy of psychology and recently in biology, by Rosenberg (1978), in that the same value of population size can be realized by a variety of underlying states or mechanisms involving lower-level variables.

[12] Blending inheritance does not require fusion of the hereditary particles. The blending refers to the character traits, not to their underlying determining factors. Unconstrained random mixing of a large number of factors having an additive effect will produce the main effect of blending, a loss of genotypic variance, until Hardy-Weinberg equilibrium is achieved. For this reason a process importantly analogous to blending in the individual case can occur even with Mendelian genetics and can be significant in the loss of genotypic variance for additive multilocus traits.

[13] These are explored in substantial further detail in the manuscript mentioned in Note 7.

[14] This conclusion must be moderated when the effect of finite population size and sampling error are taken into account. As Wade (1978, p. 110) points out, sampling error increases variance (so α would in effect be greater than 1) but at a cost of lowered group heritability. Some simple models taking account of sampling error have been constructed and will be discussed in our forthcoming paper.

[15] See Note 14.

[16] The effects are reversed for survival times of the order of one generation or less and small reproductive rates, but survival times and reproductive rates and both inclined to be appreciably larger under a wide variety of circumstances.

BIBLIOGRAPHY

Ando, A., F. M. Fisher, and H. A. Simon: 1963, *Essays on the Structure of Social Science Models*, MIT Press, Cambridge.

Bantz, D. A.: 1976, *Does Physics Explain Chemistry?*, unpublished Ph.D. Dissertation, Committee on Conceptual Foundations of Science, Univ. of Chicago.

Bogaard, P. A.: 1979, 'A Reductionist's Dilemma: the Case of Quantum Chemistry', in I. Hacking and P. D. Asquith (eds.), *PSA 1978*, Vol. 2, Philosophy of Science Association, East Lansing, Michigan.

Boyd, R.: 1972, 'Determinism, Laws and Predictability in Principle', *Philosophy of Science* 39, 431–450.

Campbell, D. T.: 1974, 'Evolutionary Epistemology', in P. A. Schilpp (ed.), *The Philosophy of Karl Popper*, Open Court, La Salle, Illinois, Vol. 1, pp. 413–463.

Campbell, D. T.: 1977, 'Descriptive Epistemology: Psychological, Sociological, and Evolutionary', The William James Lectures, Harvard, Spring, 1977, unpublished ms.
Cassidy, John: 1978, 'Philosophical Aspects of the Group Selection Controversy', *Philosophy of Science* 45; 575–594.
Causey, R. L.: 1972*a*, 'Uniform Microreductions', *Syntheses* 25, 176–218.
Causey, R. L.: 1972*b*, 'Attribute-identities in Microreductions', *Journal of Philosophy* 69, 407–422.
Causey, R. L.: 1977, *Unity of Science*, Reidel, Dordrecht, Holland.
Cohen, R. S., C. A. Hooker, A. C. Michalos, and J. van Evra (eds.): 1976, *PSA 1974* (Boston Studies in the Philosophy of Science, Vol. 32) Reidel, Dordrecht, Holland.
Darden, L., and N. Maull: 1977, 'Interfield Theories', *Philosophy of Science* 44, 43–64.
Dawkins, R.: 1976, *The Selfish Gene*, Oxford Univ. Press, New York.
Dawikins, R.: 1978, 'Replicator Selection and the Extended Phenotype', *Zeitschrift für Tierpsychologie* 47, 61–76.
Dresden, M.: 1974, 'Reflections on Fundamentality and Complexity', in C. Enz and J. Mehra (eds.), *Physical Reality and Mathematical Description*, Reidel, Dordrecht, Holland, pp. 133–166.
Fisher, R. A.: 1930, *The Genetical Theory of Natural Selection*, Oxford Univ. Press, London.
Glassman, R. B.: 1978, 'The Logic of the Lesion Experiment and Its Role in the Neural Sciences', in S. Finger (ed.), *Recovery from Brain Damage: Research and Theory*, Plenum Press, New York, pp. 3–31.
Globus, G., G. Maxwell, and I. Savodnik (eds.): 1976, *Brain and Consciousness: Scientific and Philosophic Strategies*, Plenum Press, New York.
Gregory, R. L.: 1967, 'Models and the Localization of Function in the Central Nervous System', reprinted in C. R. Evans and A. D. J. Robertson (eds.), *Key Papers: Cybernetics*, Butterworths, London, pp. 91–102.
Hull, D. L.: 1973, 'Reduction in Genetics – Biology or Philosophy?', *Philosophy of Science* 39, 491–498.
Hull, D. L.: 1974, *Philosophy of Biological Science*, Prentice-Hall, Englewood Cliffs, New Jersey.
Hull, D. L.: 1976, 'Informal Aspects of Theory Reduction', in Cohen *et al.* (1976), pp. 633–652.
Hull, D. L.: 1978*a*, 'A Matter of Individuality', *Philosophy of Science* 45, 335–360.
Hull, D. L.: 1978*b*, 'The Units of Evolution: A Metaphysical Essay', presented at the University of Aarhus symposium 'Evolution: Its Philosophical and Methodological Aspects', in August, 1978. A later version of this paper will be given at the 1980 meetings of the Philosophy of Science Association, and published in the proceedings.
Jenkin, Fleeming: 1867, 'Review of Darwin's *Origin of Species*', *The North British Review*, reprinted in David L. Hull (ed.), *Darwin and His Critics*, Harvard Univ. Press, Cambridge, 1973.
Levins, R.: 1966, 'The Strategy of Model Building in Population Biology', *American Scientist* 54, 421–431.
Levins, R.: 1968, *Evolution in Changing Environments*, Princeton Univ. Press, Princeton.
Levins, R.: 1970, 'Extinction', in *Some Mathematical Questions in Biology: lectures on mathematics in the life sciences, American Mathematical Society* 2, 75–108.

Lewontin, R. C.: 1970, 'The Units of Selection', *Annual Review of Ecology and Systematics* 1, 1–18.
Lewontin, R. C.: 1974, *The Genetic Basis of Evolutionary Change*, Columbia Univ. Press, New York.
Lewontin, R. C., and M. J. D. White: 1960, 'Interaction between inversion polymorphisms of two chromosome pairs in the grasshopper, *Moraba scurra*', *Evolution* 14, 116–129.
Lorenz, K. Z.: 1965, *Evolution and the Modification of Behavior*, Univ. of Chicago Press, Chicago.
Maull, N. L.: 1977, 'Unifying Science Without Reduction', *Studies in History and Philosophy of Science* 8, 143–162.
Maull-Roth, N.: 1974, *Progress in Modern Biology: An Alternative to Reduction*, unpublished Ph.D. Dissertation, Committee on Conceptual Foundations of Science, Univ. of Chicago.
Maynard Smith, John: 1975, *The Theory of Evolution*, 3rd ed., Penguin, London.
Maynard Smith, John: 1978, *The Evolution of Sex*, Cambridge Univ. Press, London.
McCauley, D. E., and M. J. Wade: 1979, 'Group selection: the genetic and demographic basis for the phenotypic differentiation of small populations of *Tribolium castaneum*', *Evolution* (submitted).
McCauley, Robert N.: 1979, *Explanation, Cross-Scientific Study, and the Philosophy of Mind: An Examination of the Methodological Foundations of Transformational Generative Grammar*, Ph.D. Dissertation, Department of Philosophy, Univ. of Chicago.
Moss, R. Y.: 1979, 'Designing Reliability into Electronic Components', Hewlett-Packard, Palo Alto, California.
Nagel, E.: 1961, *The Structure of Science*, Harcourt, Brace, New York.
Newell, A., and H. A. Simon: 1972, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, New Jersey.
Newell, A., J. C. Shaw, and H. A. Simon: 1957, 'Empirical Explorations with the Logic Theory Machine: A Case Study in Heuristics', reprinted in E. A. Feigenbaum and J. Feldman (eds.), *Computers and Thought*, McGraw-Hill, New York, 1963.
Newell, A., and H. A. Simon: 1961, 'GPS, a Program that Simulates Human Thought', reprinted in E. A. Feigenbaum and J. Feldman (eds.), *Computers and Thought*, McGraw-Hill, New York, 1963.
Nickles, T.: 1973, 'Two Concepts of Intertheoretic Reduction', *Journal of Philosophy* 70, 181–201.
Nickles, T.: 1976, 'Theory Generalization, Problem Reduction, and the Unity of Science', in Cohen *et al.* (1976), pp. 33–74.
Nilsson, N. J.: 1971, *Problem-Solving Methods in Artificial Intelligence*, McGraw-Hill, New York.
Rosenberg, A.: 1978, 'The Supervenience of Biological Concepts', *Philosophy of Science* 45, 368–386.
Roughgarden, Jonathan: 1979, *Theory of Population Genetics and Evolutionary Ecology: An Introduction*, Macmillan, New York.
Ruse, M.: 1973, *The Philosophy of Biology*, Hutchinson University Library, London.
Ruse, M.: 1976, 'Reduction in Genetics', in Cohen *et al.* (1976), pp. 633–651.
Salmon, W. C.: 1971, *Statistical Explanation and Statistical Relevance*, Univ. of Pittsburgh Press, Pittsburgh.

Schaffner, K. F.: 1967, 'Approaches to Reduction', *Philosophy of Science* 34, 137–147.
Schaffner, K. F.: 1969, 'The Watson-Crick Model and Reductionism', *British Journal for the Philosophy of Science* 20, 325–348.
Schaffner, K. F.: 1974, 'The Peripherality of Reductionism in the Development of Molecular Biology', *Journal for the History of Biology* 7, 111–139.
Schaffner, K. F.: 1976, 'Reduction in Biology: Prospects and Problems', in Cohen *et al.* (1976), pp. 613–632.
Schaffner, K. F.: 1978, *The Logic of Discovery and Justification in the Biomedical Sciences*, unpublished manuscript, Department of History and Philosophy of Science, Univ. of Pittsburgh.
Schaffner, K. F.: 1979, 'Theory Structure in the Biomedical Sciences', *Journal of Medicine and Philosophy* 3, 331–371.
Shimony, A.: 1970, 'Statistical Inference', in R. G. Colodny (ed.), *The Nature and Function of Scientific Theories*, Univ. of Pittsburgh Press, Pittsburgh, pp. 79–172.
Simon, H. A.: 1957, 'A Behavioral Model of Rational Choice', in H. A. Simon, *Models of Man*, Wiley, New York, pp. 241–256. See also introduction to Part IV, pp. 196–206.
Simon, H. A.: 1966*a*, 'Thinking by Computers', in R. G. Colodny (ed.), *Mind and Cosmos*, Univ. of Pittsburgh Press, Pittsburgh, pp. 3–21.
Simon, H. A.: 1966*b*, 'Scientific Discovery and the Psychology of Problem Solving', in R. G. Colodny (ed.), *Mind and Cosmos*, Univ. of Pittsburgh Press, Pittsburgh, pp. 22–40.
Simon, H. A.: 1968, *The Sciences of the Artificial*, MIT Press, Cambridge.
Simon, H. A.: 1973, 'The Structure of Ill-Structured Problems', *Artificial Intelligence* 4, 181–201.
Sklar, L.: 1973, 'Statistical Explanation and Ergodic Theory', *Philosophy of Science* 40, 194–212.
Sklar, L.: 1976, 'Thermodynamics, Statistical Mechanics, and the Complexity of Reductions', in Cohen *et al.* (1976), pp. 15–32.
Sober, E.: 1979, 'Significant Units and the Group Selection Controversy', unpublished ms., Department of Philosophy, University of Wisconsin at Madison. A later version of this paper will be given at the 1980 meetings of the Philosophy of Science Association and published with the proceedings.
Tversky, A., and D. Kahneman: 1974, 'Decision under Uncertainty: Heuristics and Biases', *Science* 185, 1124–1131.
Valenstein, E.: 1973, *Brain Control*, Wiley, New York.
Wade, Michael J.: 1976, 'Group Selection among laboratory populations of *Tribolium*', *Proceedings of the National Academy of Sciences, U.S.A.* 73, 4604–4607.
Wade, Michael J.: 1977, 'An Experimental Study of Group Selection', *Evolution* 31, 134–153.
Wade, Michael J.: 1978, 'A Critical Review of the Models of Group Selection', *Quarterly Review of Biology* 53, 101–114.
Wade, Michael J.: 1979, 'The Evolution of Social Interactions by Family Selection', *American Naturalist* 113, 399–417.
Wade, Michael J., and D. E. McCauley: 1979, 'Group Selection: The phenotypic and genotypic differentiation of small populations', *Evolution* (submitted).
Wade, Michael J., and W. C. Wimsatt: Forthcoming, 'The Blending Theory of Group Inheritance'.

Wallace, B.: 1968, *Topics in Population Genetics*, Norton, New York.
Williams, G. C.: 1966, *Adaptation and Natural Selection*, Princeton Univ. Press, Princeton.
Wilson, E. O.: 1975, *Sociobiology: The New Synthesis*, Harvard Univ. Press, Cambridge.
Wimsatt, W. C.: 1974, 'Complexity and Organization', in K. F. Schaffner and R. S. Cohen (eds.), *PSA 1972*, (Boston Studies in the Philosophy of Science, Vol. 20), Reidel, Dordrecht, Holland, pp. 67–86.
Wimsatt, W. C.: 1976*a*, 'Reductionism, Levels of Organization, and the Mind-Body Problem', in Globus *et al.* (1976), pp. 199–267.
Wimsatt, W. C.: 1976*b*, 'Reductive Explanation: A Functional Account', in Cohen *et al.* (1976), pp. 671–710.
Wimsatt, W. C.: 1978, 'Reduction and Reductionism', in H. Kyburg, Jr., and P. D. Asquith (eds.), *Current Problems in Philosophy of Science*, Philosophy of Science Association, East Lansing, Michigan.
Wimsatt, W. C.: 1980, 'Randomness and Perceived-Randomness in Evolutionary Biology', *Synthese* 43, 287–329.
Wimsatt, W. C.: 1980, 'Robustness and Functional Localization: Heuristics for Determining the Boundaries of Systems and their Biases', in M. Brewer and B. Collins (eds.), *Knowing and Validating in the Social Sciences: A Tribute to Donald T. Campbell*, Jossey-Bass, San Francisco.
Winston, P. H.: 1977, *Artificial Intelligence*, Addison-Wesley, Reading, Massachusetts.
Wright, Sewall: 1968, *Evolution and the Genetics of Populations*, Vol. 1, Univ. of Chicago Press, Chicago.